

```
Db      668  CCCACAGACTATGGAACAGACGCTCTTATGTACACAGCGCGGAGGGTTCGAC 727
Qy      121  GlyThrTyrGlnThrAsnValValArgSerGlyGlySerCysLeuTyrValProPro 140
Db      728  AGCAGCTATCCAAAGAACTGGTGGTGGGAACAACAGCGCTGCTGTGTAGTCCGCC 787
Qy      141  GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHis 160
Db      788  GGCATCTTCAAGACCACTGCAAGATCGACATCACCTGGTTCCTCCCTTCGACGACCAACA 847
Qy      161  CysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLys 180
Db      848  TCGAGATGAAGTTTGGCAGCTGACCTATGATGGTTATCAGTTGGATCTACAACATACAG 907
Qy      181  AspGluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGly 200
Db      908  GATAGAGGGGCGGAGATATAAGCAGTTTGTACGAATGGCGAATGGGAGTTAATAGGA 967
Qy      201  MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspVal 220
Db      968  GTCCCCGGCAAGCCACAGAGATCTACTACACTGTTGTTCGGAGCCATACGACATC 1027
Qy      221  ThrPheThrIleMetIleArgArgThrLeuTyrTyrPhePheAsnLeuIleValPro 240
Db      1028  ACGTTTCGGTGGTGATCCGGAGGAAACGCTCTACTACTTCTTCAATCTGATCGTSCCC 1087
Qy      241  CysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 260
Db      1088  TGGTGTCTCATCGCCTCCATGGCTCTATTGGGGTTCACCTTGCTCCAGACTCCGGAGAA 1147
Qy      261  LysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAla 280
Db      1148  AAGTTGTCTTAGGTGACGATATTAAGTGTGCTGTTGACGGTGTCTTCAACATGGTGCGC 1207
Qy      281  GluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrTyrPheAsnCysIle 300
Db      1208  GAGACGATGCCAGCGACGTCGGACCGCTGCGCCCTGCTCGGCACCTACTTCAACTGCATC 1267
Qy      301  MetPheMetValAlaSerSerValValLeuThrValValValLeuAsnTyrHisArg 320
Db      1268  ATGTTCTATGTGTGCTTCTCCGCTGCTCCACCATACTGATCTCTCAACTACCACCGG 1327
Qy      321  ThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrpLeuPro 340
Db      1328  CAGCAGACACTCAGCAATAGAGTAGTATGATTCGTTGCTGTTCTTATTTGCTGCCG 1387
Qy      341  TrpIleLeuArgMetSerArgProGlyLysIleThrArgLysThrIleMetMetAsn 360
Db      1388  TGGTGTCTGCGCATGTCAACGCGCGGCTCGCGCAGACGCGCGCGCGCGGTA-- 1444
Qy      361  ThrArgMetArgGluLeuGluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnVal 380
Db      1445  CTTCCGCGCGCGGACCTGGAGCTGCGGAGCGCTCTCCCAAGTGGCTCTTAGGCAAGTG 1504
Qy      381  LeuAspIleAspAspPheArgHisGlyProProProProAsnSerThrAlaSerThr 400
Db      1505  CTGCAGACTCATGACGACTTCCGCCAC-----CCGCAAGCGCAGCAG----- 1546
Qy      401  GlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgArgSerPheValArg 420
Db      1547  -----CCGCAATGCTGCCGATAC-----TACAGGGGGGGT----- 1576
Qy      421  ProSerThrMetGluAspValGlyGlyLeuGlySerHis-----His 435
Db      1577  -----GAGGAGAAATGGCGGGGTGGCGCGCGCACAGTTGCTTTCGGTGTTCGAC 1624
Qy      436  ArgGluLeuHisLeuIleLeuArgGluLeuPheIleThrAlaArgMetLysLysAla 455
Db      1625  TACGAGCTCTCCCTCATCTTGAAGAGATATAGATCATCAGATCAGATGCGCAGGAC 1684
Qy      456  AspGluGluAlaGluLeuIleSerAspTrpLysPheAlaAlaMetValValAspArgPhe 475
Db      1685  GACGAGATCGGACATTTCCGCGGCACTGGAGTTCCGCCGCACTGGTGTGGACAGACTG 1744
```

```
Qy      476  CysLeuPheValPheThrLeuPheThrIleIleAlaThrValAlaValLeuLeuSerAla 495
Db      1745  TGCCTTATTATCTTACCTGTTTCACAATCATGCCACGCTAGCGTGCTGCTCGCGC 1804
Qy      496  ProHisIleIleVal 500
Db      1805  CCACACATCATGGTG 1819

RESULT 12
AX009612          3700 bp      DNA      linear      PAT 06-SEP-2000
LOCUS            Sequence 3 from Patent EP0962528.
DEFINITION       AX009612
ACCESSION       AX009612
VERSION         AX009612.1 GI:9996844
KEYWORDS
SOURCE          Heliothis virescens (tobacco budworm)
ORGANISM        Heliothis virescens
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
                Noctuoidea; Noctuidae; Heliothinae; Heliothis.

REFERENCE
AUTHORS         Adamczewski,M.D., Schulte,T.D. and Oellers,N.D.
TITLE           Nucleic acids encoding acetylcholin-receptor subunits from insects
JOURNAL         Patent: EP 0962528-A 3 08-DEC-1999;
                BAYER AG (DE)

FEATURES         Location/Qualifiers
source           1..3700
                 /organism="Heliothis virescens"
                 /mol_type="unassigned DNA"
                 /db_xref="taxon:7102"
                 335..1825
                 /note="unnamed protein product"
                 /codon_start=1
                 /protein_id="CAC07500.1"
                 /db_xref="GI:9996845"
                 /db_xref="RENTREMBL:CAC07500"
                 /translation="MGRARRSLAAPAGLLLLCLLWPRGARGVHEKRLHLLHLDH
                 YNVLRFVNESDPLQSLFGLTQIIVDEKNQLITNWLKLEWMDNMLRWNTSDF
                 GVKDLRWPFPHLWKPDLVMSADGDFSTYPTNVVRNNGSLVYPGIFKSTCKI
                 DIWFPPDQRCMEKFGSTYDGYQLDLOLDQGGDGISSFVTNGEWELGVPKRNE
                 IYVCCPEYDITFAVVRRTLYVFENLIVPCVLASMAILLGFTLPPDSGEKLSIG
                 VTILLSTVFLNVAETPATSDAVELLGTTFYFNCIMFVWASSVSTILLNLYHRRHAD
                 THEMSDWIKCVFLWLPVWRMSRPGSATTPPARPPPPDLELRSSKSLIANVLD
                 IDDFRHPQAQQPCQYRYRGEENAGLAHSCFGVDYELSLILKEIRVITDQMRKD
                 DEDADISRDKWFAAMVVDRLLCIIFTLTIIATLAVLLSAPHINVS"

ORIGIN
Alignment Scores:
Pred. No.:       7,27e-164      Length:      3700
Score:           1803.50        Matches:     347
Percent Similarity: 79.21%      Conservative: 53
Best Local Similarity: 68.71%   Mismatches:  78
Query Match:     68.31%        Indels:      27
DB:              6              Gaps:        7

US-09-303-232-6 (1-501) x AX009612 (1-3700)

Qy      2   AlaProMetLeuAlaAlaLeuAlaLeuLeuAlaLeuProValSerGluGln--Gly 20
Db      368  GCGCCCGCGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 427
Qy      21  ProHisGluLysArgLeuLeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgPro 40
Db      428  TACCACGAGAGCGGCTACTGTCACCACCTATTGGACCACTTACCAACGTACTGGAGAGGCC 487
Qy      41  ValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleIle 60
Db      488  GTCGTCAACGAGAGACCGCGCTGCGCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 547
Qy      61  AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsn 80
Db      548  GACGTGGAAGAGNAAGAACGAGCTTTTATTAACAACATCTGGCTAAACTAGAGTGAAT 607
```


[illegible]

Pred. No.:	3.99e-164	Length:	1683
Score:	1801.50	Matches:	347
Percent Similarity:	76.60%	Conservative:	59
Best Local Similarity:	65.47%	Mismatches:	79
Query Match:	68.24%	Indels:	45
DB:	3	Gaps:	5

US-09-303-232-6 (1-501) x DME554210 (1-1683)

Qy	4	MetLeuAlaLaLeuAlaLeuLeuAlaLeuLeuProValSerGluGlnGlyProHisGlu	23
Db	83	CTGCTGCTATGGCTGGGACTTTTAAATTTATGATACCGCTTTGTGGCTGACCCCATGAG	142
Qy	24	LysArgLeuLeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsn	43
Db	143	AAGCGGCTACTCCACGCCCTTCTGGACAACCTACCAACAGCTGGAGCGTCGGTGGTCAAT	202
Qy	44	GluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleLeuAspValAsp	63
Db	203	GAATCCGATCCATTGCCAATCGAGCTTCGGACTTAACACTCATGTCAGATTATCGATGTGGAC	262
Qy	64	GluLysAsnGlnLeuLeuIleThrAsnIleThrLeuSerLeuGluThrAsnAspTyrAsn	83
Db	263	GAAGAAGATCAACTGCTTATAACGAATATTGGCTCAAAATGGGAATGGAAACGATATGAAT	322
Qy	84	LeuArgTrpAsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsnLys	103
Db	323	CTTCGATGGGAATTCGAGTGAGTTCCGGTGGTGGGGATCTCGGAAATCCGCCACATATCGC	382
Qy	104	LeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyr	123
Db	383	CTATGGAAACCGGATGTACTGATGTACAACAGTCCCGACGAGGCTTCGATGGAAACGTAC	442
Qy	124	GlnThrAsnValValValArgSerGlyGlySerCysLeuTyrValProGlyIlePhe	143
Db	443	GCCACAAATGTGGTGGTTCGCAATTAATGGGAGCTGTCTGTACGTACCGCCAGGTATATTT	502
Qy	144	LysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMet	163
Db	503	AAGTCAACGTGAAGATCGACATTACGTGGTTCCTATTCGACGATCAGAGATGTGAATG	562
Qy	164	LysPheGlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAla	183
Db	563	AAATTTGGTTCGTGGACCTACGATGGGTTCAGTTGGACCTGCAGTTGCAGGACGAAGCT	622
Qy	184	GlyGlyAspLeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGlyMetProGly	203
Db	623	GGTGGCGACATTTCTAGCTTTATTAACAAATGGCAATGGGACTTTGTTAGGTGTGCCCGT	682
Qy	204	LysLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThr	223
Db	693	AAACGAAATGAATCTACTATAATGTCTGCCACAGCTTATATTGACATTAACAAATCGCC	742
Qy	224	IleMetIleArgArgTrpLeuTyrTyrPhePheAsnLeuIleValProCysValLeu	243
Db	743	ATTTTGATAGGCGCAAAACGTGTACTATTTTTTCAATCTGATGTGGCGGTGGTACTG	802
Qy	244	IleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThr	263
Db	803	ATCGCTCCATGGCACTGCTAGGGTTTACACTGTCACAGATTCTGGTGAAGAGCTTTTCG	862
Qy	264	LeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeu	283
Db	863	CTTGAGGTATCAATTTCTATTATCGCTTACAGTCTTCTTCAACATGTGGTGGCGCAACATG	922
Qy	284	ProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMet	303
Db	923	CCGGCGACCTCCGATCGGTACCGCTGCTCGTAAAGTATTTCAATTCGATTAATGTTTATG	982
Qy	304	ValAlaSerSerValValLeuThrValValValLeuAsnTyrHisArgThrAlaAsp	323
Db	993	GTGGGCTCATAGTTGTGTCAACACATATCTGCTCAATTCATCATATAGAAATCCAGAT	1042

Qy	324	Il eH i sGluMetProGlnTrpLl eLysSerValPheLeuGlnTrpLeuProTrpLl eLeu	343
Db	1043	ACGCATGAATGAGTGAATCGATAAGAGTAATAATTCCTTTATTGGTTACCTTGCATATTTG	1102
Qy	344	ArgMetSerArgProGlyLys-----Lyslle	352
Db	1103	CGCATGCAAGACCCGGACAGTGTGGCTACGAATGCCCGCGCGCCTCTTCTTCGAGT	1162
Qy	353	ThrArgLysThrIleMetMetAsnThrArgMetArgLuluLeuGluLysGluArgSer	372
Db	1163	TCCTCCGCATCCGGCGAGAGAAGACACATCCAAACCGTTTCAGGTTCAGGAGAGGTC	1222
Qy	373	SerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisGlyProPro	392
Db	1223	TCCAAGTCTCTGCTGGCCAAATGCTCGATATAGACGATATTCCGA-----	1270
Qy	393	ProProAsnSerThrAlaSerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThr	412
Db	1271	--TGCAATCATCGATGCGCAGCGGACCTTTGCCCCACCAGGCCACATATTACAGGACG	1327
Qy	413	AspPheArgArgSerPheValArgProSerThrMetGluAsp-----	426
Db	1328	ATGTACAGGCAAGGG-----GATACGGCAGCGTGGGACCCCGT	1366
Qy	427	-----ValGlyGlyGlyLeu-----	431
Db	1367	GGACCACCTGGTCCAGTTGTGGACGGCGTGTGCACGAGGCCATTTCCACACCTGTCGT	1426
Qy	432	---GlySerHisHisArgGluLeuHisLeuIleLeuArgGluLeuGlnPheIleThrAla	450
Db	1427	ACATCTCTCGGAGTAGTACGAACCTGGCCCTCATACTCAAGGAGCTGCGTTGGAATAACGAA	1486
Qy	451	ArgMetLysLysAlaAspGluGluAlaGluLeuIleSerAspTrpLysPheAlaAlaMet	470
Db	1487	CAGCTCAAAAAGAGGACGAAACAGGACGACATTACGCGAGATTGGAATTTGCTGCCATG	1546
Qy	471	ValValAspArgPheCysLeuPheValPheThrLeuPheThrIleLeuAlaThrValAla	490
Db	1547	GTCGTCGATCGTTGTGCTCTATTATTTTTCACCTTGTATTACTATTATAGCAACCTCGCT	1606
Qy	491	ValLeuLeuSerAlaProHisIleIleVal	500
Db	1607	GTACTCTTCAGCGCCGCAATTCATTGTG	1636
RESULT 15			
LOCUS	AF272778	2907 bp mRNA linear	INV 29-APR-2002
DEFINITION	Drosophila melanogaster nicotinic acetylcholine receptor Dalphas5 subunit (nACKalpha-34E) mRNA, nACKalpha-34E-A allele, complete cds.		
ACCESSION	AF272778		
VERSION	AF272778.1	GI:20152839	
KEYWORDS			
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 2907)		
AUTHORS	Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.		
TITLE	Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing		
JOURNAL	Genetics 160 (4), 1519-1533 (2002)		
MEDLINE	21969411		
PUBMED	11973307		
REFERENCE	2 (bases 1 to 2907)		
AUTHORS	Grauso, M. and Sattelle, D.B.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-MAY-2000) Human Anatomy and Genetics, MRC-FGU, University of Oxford, South Park Road, Oxford OX1 3QK, UK		
FEATURES	Location/Qualifiers		
source	1. .2907		

gene	/organism="Drosophila melanogaster"	Best Local Similarity: 66.29%	Mismatches: 69
	/mol_type="mRNA"	Query Match: 68.20%	Indels: 47
	/db_xref="taxon:7227"	DB: 3	Gaps: 9
	/chromosome="2"		
	/map="2L: 34E4-5"		
	/dev_stage="embryo"		
1. .2907		US-09-303-232-6 (1-501) x AP272778 (1-2907)	
CDS	/gene="nAcRalpha-34E"		
	/allele="A"		
289. .2712			
	/note="ion channel; neurotransmitter transmembrane receptor"		
	/codon_start=1		
	/product="nicotinic acetylcholine receptor Dalpha5 subunit"		
	/protein_id="AAMI3390.1"		
	/db_xref="GI:20152840"		
	/translation="MKNAQLKLTVEVDDDELWLAVALHSCSNFSSSSSTRITSSNORH NQQLTTLQPSLSIKKHSNIASEOHNSQQQEPASKDEEDVANHGSRNDQTHLQOLDSS NMLSPKTAATAAGTAGDEATTCQPTNIRLCARKQRLRRKRRKRPATNETDIIKQQQL SMPEFKRKSTDTISPAATTCPTATMQCRASDNEFSPI SKHDRVSTATFAWLH VLQVLLSQOQLHVQORSVLLFRITAASTIAFISYLGFAAQLKNSSSSSSSNS NNGSTQLNGINKHSWIFLYLNLAKVCLAGHEKRLHLDLPYNTLRRPVNES DPQLSFLGLTMOIIVDEKNQLIVTWLKLWMDMLRWNTSDYGVKDLRIPRER IKPDLVYNASADEGFGDTQTNVVRNNGSCLYVPPGIPKSKTIDITWEPDPDQRC EMKGSWTYDGFOLDLQDETGDSSIVNGEWELLGVPGKRNIEIYNCCPEPYID ITAIIRRLIYFFNLIIPCVLIASMALLGFTLPDPSGEKLSGLVTILLISUTVFIN MVAETPATSADVPLLTGFNCIMFVASSVSTILLNHYHRNADTHEMSWIRIVF LCWPLTILMRSGRPLILEFPPTPCSDTSERKQLISDVKELRSKSKILLANLVDI DDPFRNCRMTPGTILPHNPAFVRTVYGGDDGSGIPIGSTRMPDAVTHHTCIKSST EYELGLTLEIRFITDQLRKDECDNDIANDKFAAMVVDRLCLLIIFTFMFAITAVL LSAPHIIVS"		
variation	/gene="nAcRalpha-34E"		
	/note="compared to B allele"		
375	/replace="a"		
variation	/gene="nAcRalpha-34E"		
	/note="results in phenylalanine to isoleucine substitution; compared to B allele"		
827	/replace="c"		
variation	/gene="nAcRalpha-34E"		
	/note="results in threonine to isoleucine substitution; compared to B allele"		
1079. .1080	/replace="t"		
variation	/gene="nAcRalpha-34E"		
	/note="results in lysine to arginine substitution; compared to B allele"		
1081. .1086	/replace="gg"		
variation	/gene="nAcRalpha-34E"		
	/note="results in deletion of asparagine and serine; compared to B allele"		
1119. .1120	/replace=""		
variation	/gene="nAcRalpha-34E"		
	/note="results in serine, glycine, and asparagine insertion; compared to B allele"		
1121	/replace="agcggcaac"		
variation	/gene="nAcRalpha-34E"		
	/note="results in asparagine to serine substitution; compared to B allele"		
	/replace="g"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.02e-163	Length: 2907	
Score:	1800.50	Matches: 348	
Percent Similarity:	77.90%	Conservative: 61	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 06:39:04 ; Search time 526.234 Seconds

(without alignments)
4044.486 Million cell updates/sec

Title: US-09-303-232-6

Perfect score: 2640

Sequence: 1 MAPMLAALALLPVSEQ.....LFTIATVALLSAPHIIVQ 501

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh

-Q=/cgn2_1/USPTO.spool/US09303232/runat_07052004_101109_23873/app_query.fasta_1.2261

-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blastm62 -TRANS=human40.cdi

-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09303232.cgn_1_1200 @runat_07052004_101109_23873 -NCFU=6 -ICPU=3

-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2640	100.0	3109	3 AAZ24477	Aaz24477 H. viresc
2	1968	74.5	1540	4 ABL07231	Abi07231 Drosophil
3	1803.5	68.3	3700	3 AAZ24476	Aaz24476 H. viresc
4	1609	60.9	2886	3 AAZ24475	Aaz24475 D. melano
5	1283	48.6	936	4 ABL13733	Abi13733 Drosophil
6	1258.5	47.7	1509	3 AAC58395	Aac58395 Human PRO
7	1258.5	47.7	1509	4 AAC90380	Aac90380 Wild-type
8	1258.5	47.7	1876	2 AAV12197	Aav12197 Human neu

9	1258.5	47.7	1876	2	AAZ48239	Aat48239 Neuronal
10	1258.5	47.7	1876	6	ABS54875	AbS54875 Human neu
11	1258.5	47.7	1876	6	ABV73248	Abv73248 Human neu
12	1258.5	47.7	1876	8	ADA10864	Ada10864 Human neu
13	1254.5	47.5	1590	2	AAV44687	Aav44687 V27AT var
14	1253	47.5	2769	2	AAT59196	Aat59196 Neuronal
15	1252.5	47.4	1509	4	AAC90385	Aac90385 Mutant hu
16	1248.5	47.3	1509	4	AAC90386	Aac90386 Mutant hu
17	1242.5	47.1	1509	4	AAC90387	Aac90387 Mutant hu
18	1229	46.6	1964	6	ABZ11298	Abz11298 Human pol
19	1226.5	46.5	2101	2	AAT59197	Aat59197 Neuronal
20	1080	40.9	1915	4	AAD20962	Aad20962 Caenorhab
21	1046	39.6	803	4	ABL07799	Abi07799 Drosophil
22	989.5	37.5	1908	2	AAT48236	Aat48236 Neuronal
23	989.5	37.5	1908	8	ADA10856	Ada10856 Human neu
24	982.5	37.2	1584	7	ADA83809	Ada83809 Human CHR
25	979	37.1	1416	6	AAC90382	Aac90382 Chimeric
26	946.5	35.9	1654	6	ABS54871	AbS54871 Human neu
27	946.5	35.9	1654	6	ABV73244	Abv73244 Human neu
28	943	35.7	2210	4	ABL10275	Abi10275 Drosophil
29	940.5	35.6	1757	2	AAV12200	Aav12200 Human neu
30	933	35.3	1743	2	AAT48232	Aat48232 Neuronal
31	933	35.3	1743	6	ABS54874	AbS54874 Human neu
32	933	35.3	1743	6	ABV73247	Abv73247 Human neu
33	933	35.3	1743	8	ADA10862	Ada10862 Human neu
34	933	35.3	2385	4	ABL11821	Abi11821 Drosophil
35	922.5	34.9	2277	2	AAV12199	Aav12199 Human neu
36	922.5	34.9	2277	2	AAQ90387	Aaq90387 Alpha 2 s
37	922.5	34.9	2277	6	ABS54870	AbS54870 Human neu
38	922.5	34.9	2277	6	ABV73243	Abv73243 Human neu
39	922.5	34.9	2664	2	AAT48235	Aat48235 Neuronal
40	922.5	34.9	2664	6	ABK92165	Abk92165 Prostate
41	922.5	34.9	2664	8	ADA10854	Ada10854 Human neu
42	922.5	34.9	2664	9	ADC71170	Adc71170 Human 205
43	922.5	34.9	2666	9	ADC71168	Adc71168 Human 205
44	922	34.9	1503	4	AAD20961	Aad20961 Caenorhab
45	918.5	34.8	1896	6	AAL45867	Aal45867 Modified

ALIGNMENTS

RESULT 1

AAZ24477

ID AAZ24477 standard; cDNA to mRNA; 3109 BP.

XX AC AAZ24477;

XX AC AAZ24477;

DT 17-FEB-2000 (first entry)

XX H. virescens acetyl-choline receptor DNA from clone Hva7-2.

DE H. virescens acetyl-choline receptor DNA from clone Hva7-2.

XX Acetyl-choline receptor; nicotinic; insect; insecticide; screening;

KW neurotransmission; plant protection agent; conductance; AChR; ds.

XX Heliothis virescens.

XX Heliothis virescens.

XX DE19819829-A1.

XX 11-NOV-1999.

XX 04-MAY-1998; 98DE-01019829.

XX 04-MAY-1998; 98DE-01019829.

XX (FARB) BAYER AG.

XX Adamczewski M, Oellers N, Schulte T;

XX WPI; 2000-014207/02.

XX P-PSDB; AAV50816.

XX New nucleic acid encoding a nicotinic acetylcholine receptor from

PT insects, used to identify potential insecticides.

QY 401 GlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgArgSerPheValArg 420
 Db 1547 -----CCGCAATGCTCCGATAC-----TACAGGGGGGT----- 1576
 QY 421 ProSerThrMetGluAspValGlyGlyLeuGlySerHis-----His 435
 Db 1577 -----GAGGAGATGGCGGGGTGGCGGCGCACAGTTGCTTCGGTGTGCAC 1624
 QY 436 ArgGluLeuHisLeuLeuLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAla 455
 Db 1625 TACAGCTCTCCCTCAITCTGAGGAGATTAGAGTCATCACAGATCAGATCGCGCAAGAC 1684
 QY 456 AspGluGluAlaGluLeuLeuSerAspTrpLysPheAlaAlaMetValValAspArgPhe 475
 Db 1685 GACGAGATCGGACATTCGCGGCACTGGAAGTTGCGCGCCATGCTGCTGCGGCG 1744
 QY 476 CysLeuPheValPheThrLeuPheThrIleIleAlaThrValAlaValLeuLeuSerAla 495
 Db 1745 TGCTTATTATCTTACCTGTTTCCATCATCATGCCACGCTAGCGGTGCTGCTGCGG 1804
 QY 496 ProHisIleIleVal 500
 Db 1805 CCACACATCATGGTG 1819

RESULT 4

AAZ24475
 ID AAZ24475 standard; cDNA to mRNA; 2886 BP.

XX AC AAZ24475;
 DT 17-FEB-2000 (first entry)

XX D. melanogaster acetyl-choline receptor DNA from clone Da7.
 XX Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
 XX neurotransmission; plant protection agent; conductance; AChR; ds.
 XX Drosophila melanogaster.

XX Key Location/Qualifiers
 FT CDS 372..2684
 FT /*tag= a
 FT /product= "acetyl choline receptor."
 XX DE19819829-Al.
 XX 11-NOV-1999.

XX 04-MAY-1998; 98DE-01019829.
 XX 04-MAY-1998; 98DE-01019829.

XX (FARB) BAYER AG.
 XX Adamczewski M, Oellers N, Schulte T,
 XX WPI: 2000-014207/02.
 XX P-PSDB; AAY50814.

XX New nucleic acid encoding a nicotinic acetylcholine receptor from
 XX insects, used to identify potential insecticides.

XX Claim 1a; Page 8-12; 26pp; German.

XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
 XX acetyl-choline receptor (I) from insects which can be used as an
 XX insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
 XX (also vectors containing it, its regulatory regions, and antibodies
 XX directed against (I)-encoded proteins) are used to screen for: (a) plant
 XX protection agents that alter conductance of AChR, potentially useful as
 XX insecticides, or (b) genes which encode polypeptides that are involved in
 XX formation of functionally related AChR in insects. (I) are also used to

CC isolate and characterize the specified regulatory regions and for
 CC recombinant production of (II). This sequence encodes an acetyl-choline
 CC receptor isolated from *Drosophila melanogaster*

XX
 SQ Sequence 2886 BP; 942 A; 687 C; 640 G; 617 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.51e-162 Length: 2886
 Score: 1609.00 Matches: 319
 Percent Similarity: 71.62% Conservative: 57
 Best Local Similarity: 60.76% Mismatches: 65
 Query Match: 60.95% Indels: 84
 DB: 3 Gaps: 10

US-09-303-232-6 (1-501) x AAZ24475 (1-2886)

QY 10 LeuLeuAlaLeuLeuProValSerGluGln-----GlyProHisGluLysArg 25
 Db 1254 TTATTGATATATTGAATTATCTGCTAAAGTTGCTAGCAGGATATCATGAAAGAGA 1313
 QY 26 LeuLeuAsnAlaLeuLeuAlaAsnThrLeuGluArgProValAlaAsnGluSer 45
 Db 1314 CTGTTACAGCATCTTTGGATCCTTATAATACACTAGAACGTCCTCAATGAATCG 1373
 QY 46 GluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleIleAspValAspGluLys 65
 Db 1374 GACCCGTTACAATTAAAGCTTTGGTTAACTTTAATGCMAATTATCGATGTGGACGAGAA 1433
 QY 66 AsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAspTyrAsnLysArg 85
 Db 1434 AATCAATTGCTAGTCACTAATGTGTGTTAAACTGGAGTGGAGACGATGAATCTCCGC 1493
 QY 86 TrpAsnAspSerGluTyrGlyValLysAspLeuArgIleThrProAsnLysLeuTrp 105
 Db 1494 TGGAAACACCTCCGACTATGCGGAGTTAAGGATCTGCGAATACGCGCATCGCATCTGG 1553
 QY 106 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrThrGlnThr 125
 Db 1554 AAGCCGGACGTGCTGATGTATACAAACAGTCGGGATTTGACGACCTACCGACG 1613
 QY 126 AsnValValValArgSerGlyGlySerCysLeuTyrValProProGlyIlePheLysSer 145
 Db 1614 AAGTGGTGGTGGGAAACACGCTGCTGTATAGTTCCGCGGGGATCTTCAAGTCG 1673
 QY 146 ThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMetLysPhe 165
 Db 1674 ACGTCAAGATCGACATCACGTGGTCCCTTCGATGACCGCGGTGCGAGATGAAGTTC 1733
 QY 166 GlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyGly 185
 Db 1734 GGCAGTTGGACCTACGACGATTTCCAGCTGGATTTACAATTAAGAATGAACCTGGCGGT 1793
 QY 186 AspLeuSerAspPheIleThrAsnGlyGluTyrTrpLeuIleGlyMetProGlyLysLys 205
 Db 1794 GATATCAGCAGTTACGTGCTCAACGGCGAGTGGAACTACTGGGTGTCGCCGCAACGT 1853
 QY 206 AsnThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMet 225
 Db 1854 AACGAGATCTATTACAACCTGCTGCCCGGAACCTATATAGACATCACTTCGCCCATCATC 1913
 QY 226 IleArgArgArgThrLeuTyrTyrPheAsnLeuIleValProCysValLeuIleSer 245
 Db 1914 ATCCGCCGCGAACAACACTGTACTATTCTTCAACCTGATCATCACTTGTGTACTGTGCC 1973
 QY 246 SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThrLeuGly 265
 Db 1974 TCCATGGCCCTTGTCTGGATTCACCTGCGCCGAGATTCGGGTGAAATAATTATCGTGGGT 2033
 QY 266 ValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGln 285
 Db 2034 GTTACCATCTTGCTCTCGTGGACCGGTGTTCTGAATATGTTGTCGCGAGACAATGCCGCT 2093
 QY 286 ValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAla 305


```
Db 2094 ACTTCGATCGGTGCATTG----- 2114
|||||
Qy 306 SerSerValValLeuThrValValLeuAsnThrHisArgThrAlaAspIleHis 325
Db 2114 ----- 2114
Qy 326 GluMetProGlnTrpIleYsSerValPheLeuGlnTrpLeuProTrpIleLeuAtgMet 345
|||||
Db 2115 -----TGGATACGATCGTGTGTTTGGCTGGCTGCCATGATATCGAATG 2162
Qy 346 SerArgProGlyYsIleThrArgLys-----ThrIleMetMetAsnThrArg 362
|||||
Db 2163 AGTCGCCAGGACGACGCTGATCTAGAGTTCGCCAGCACGCCCTGTTCGGACACATCC 2222
Qy 363 -----MetArgGluLeuGluLeuLysGluArgSerSerLysSer 375
2223 TCCGAGCGGAAGCACCATGATCTCCGACGTTGAGCTGAAGAGCGCTCGTGAATCG 2282
Qy 376 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisGly-----ProProPro 393
2283 CTGCTGCCAACGTAAGTACATGATGATGACTTCCGGACCAATTGTCCGCCCATGACG 2342
Qy 394 ProAsnSerThrAlaSerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThrAsp 413
2343 CCGGCG-----GGAACACTGCCACACACCGCGCTTCTATCGCACGTT 2387
Qy 414 PheArgSerPheValArgProSerThrMetGluAspValGly-----GlyGlyLeu 431
2388 TATGGACAAGGC-----GACGATGGCAGCATTTGGGCCAATT 2423
Qy 432 GlySer-----HisHisArg----- 436
2424 GGCAGCACCGAATGCCGATGCCGTCACCCATCATAGTCATCAATCACTCACTGAA 2483
Qy 437 ---GluLeuHisLeuIleLeuArgGluLeuGlnPheIleThrAlaArgMetLysAla 455
2484 TATGAATTAGTTTAACTTAAAGGAAATTCGCTTTATACTGATCAGTACGTAAAGAT 2543
Qy 456 AspGluGluAlaGluLeuIleSerAspTrpLysPheAlaAlaMetValValAspArgPhe 475
2544 GACGAGTGCATGACATTCGCAATGGAATTTGGAATTTGCAGCTATGGTGTTCAGACAG 2603
Qy 476 CysLeuPheValPheThrLeuPheThrIleIleAlaThrValAlaValLeuLeuSerAla 495
2604 TGCCTTATCATATTCAATGTTTCGCAATATTAGCCCAATAGCTGTACTACTATCGGCA 2663
Qy 496 ProHisIleIleVal 500
2664 CCACATATTATTGTC 2678
Db 2678
RESULT 5
ABL13733
ID ABL13733 standard; cDNA; 936 BP.
AC
AC ABL13733;
DT
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 35681.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
(PKKE ) PE CORP NY.
Venter JC, Adams M, Li PWD, Myers EW;
P-PSDB; ABB69630.
WPI; 2001-656860/75.
New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signaling and cell-cell
interactions.
Claim 1; SEQ ID NO 35681; 21pp + Sequence Listing; English.
The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
sequences (AB101840-AB116175) and the encoded proteins (ABBS7737-
ABBS7702). The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 936 BP; 238 A; 213 C; 239 G; 246 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.67e-127 Length: 936
Score: 1283.00 Matches: 231
Percent Similarity: 89.69% Conservative: 30
Best Local Similarity: 79.38% Mismatches: 30
Query Match: 48.60% Indels: 0
DB: 4 Gaps: 0
US-09-303-232-6 (1-501) x ABL13733 (1-936)
Qy 4 MetLeuAlaAlaLeuAlaLeuLeuAlaLeuProValSerGluGlnGlyProHisGlu 23
Db 58 CTGCTCATGGCTGGGACCTTTTAATTATGATACCGCTTGTGGCGTGGACCCCATGAG 117
Qy 24 LysArgLeuLeuAsnAlaLeuLeuAlaAsnTrpAsnThrLeuGluArgProValAlaAsn 43
Db 118 AAGCGGCTACTCCACGCCCTTCTGGACAACTACACAGCCTGGAGCGTCCGGTGCAT 177
Qy 44 GluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleAspValAsp 63
Db 178 GAATCCGATCCATTGCAACAGTCTCGGACTAACACTCATGCAGATTATCGATGGAC 237
Qy 64 GluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAspTrpAsn 83
Db 238 GAAAGAAATCACTGCTTATAACGAATATTGGCTCAAAATGGAAATGGAACGATATGAT 297
Qy 84 LeuArgTrpAsnAspSerGluTrpGlyGlyValLysAspLeuArgIleThrProAsnLys 103
Db 298 CTTTCATGGAATTCGAGTGAAGTTCGCTGGTGTGGGATCTGCGAATTCGCCCATCGC 357
Qy 104 LeuTrpLysProAspValLeuMetTrpAsnSerAlaAspGluGlyPheAspGlyThrTyr 123
Db 358 CTATGGAAACCGGATGACTATGATACACAGTCCGACGAGGCTTCGATGGAACGTAC 417
Qy 124 GlnThrAsnValValArgSerGlyGlySerCysLeuTrpValProGlyIlePhe 143
Db 418 GCCACAAATGTGGTGTTCGCAATAATGGGAGCTGTCTGTACGTACGCCAGGTATATT 477
Qy 144 LysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMet 163
Db 478 AAGTCAACGCTGAAGATCGACATTCATCGTGGTTTCCATTCGACATCAGAGATGTAATG 537
Qy 164 LysPheGlySerTrpThrTrpAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAla 183
Db 538 AAATTTGGTTCGTGACCTACATGAGGTTTCAGTTGGACCTGCAGCTTCAGAGCAAGCT 597
```

QY 184 GlyGlyAspLeuSerAspPheLeuThrAsnGlyGluThrPheLeuLeuGlyMetProGly 203
 Db 598 GGTGGCAGCAATCTTACGTTTATACCAATGGCAATGGGACTTGTAGGTGTCCCGGT 657
 QY 204 LysLysAsnThrLeuThrTyrAlaCysCysProGluProTyrValAspValThrPheThr 223
 Db 658 AACGAAATGAAATCTACTATAATGTGTCGCCAGAACCTTATATTGACATAACATTGCCC 717
 QY 224 IleMetIleArgArgArgThrLeuTyrTyrPhePheAsnLeuLeuValProCysValLeu 243
 Db 718 ATTTTGATAAGGCCAAACGTTGTACTATTTTCAATCTGATTGTGCCGTGCGTACTG 777
 QY 244 IleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThr 263
 Db 778 ATCCCTCCATGGCACTGTAGGGTTTACAGTGCACAGATCTGTGTGAAGAGTTTCG 837
 QY 264 LeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeu 283
 Db 838 CTTGAGTTACAAATCTTATTATCGTTACAGTCTTCTCAACATGGTGGCGGAACAATG 897
 QY 284 ProGlnValSerAspAlaIleProLeuLeuGly 294
 Db 898 CCGCGCACTCCGATCGGTACCGCTGCTCGGT 930

RESULT 6

AAC58395
 ID AAC58395 standard; cDNA; 1509 BP.
 AC AAC58395;
 XX
 XX 29-JAN-2001 (first entry)
 XX Human PRO2145 nucleotide sequence SEQ ID NO:76.
 XX

Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;
 proliferation; tumorigenesis; identification; cancer; cytostatic;
 KW neurotropic; neuroprotective; antiinflammatory; immunosuppressive;
 KW immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy;
 KW neuronal disorder; glial disorder; astrocytal disorder; angiogenic;
 KW hypothalamic disorder; glandular disorder; macrophagal disorder;
 KW epithelial disorder; stromal disorder; blastocoelec disorder;
 KW inflammatory disorder; immunologic disorder; ss.
 XX Homo sapiens.
 OS
 XX WO200053755-A2.
 FN
 XX 14-SEP-2000.
 PD
 XX 06-JAN-2000; 2000WO-US000376.
 XX
 XX 08-MAR-1999; 99WO-US0005028.
 PR
 XX 02-JUN-1999; 99WO-US012252.
 PR
 XX 23-JUN-1999; 99US-0141037P.
 PR
 XX 07-JUL-1999; 99US-0143048P.
 PR
 XX 26-JUL-1999; 99US-0145698P.
 PR
 XX 30-NOV-1999; 99WO-US028313.
 PR
 XX 20-DEC-1999; 99WO-US030911.
 PR
 XX 05-JAN-2000; 2000WO-US000219.
 XX

(GETH) GENENTECH INC.

PA Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA;
 PI Watanabe CK, Wood WI;
 PI WPI; 2000-572270/53.
 DR P-PSDB; AAB24088.
 DR
 XX Thirty PRO polynucleotides encoding PRO polypeptides, useful in the
 PT treatment, diagnosis and prevention of cancer.
 PT
 XX Claim 50; Fig 57; 286pp; English.
 PS

XX

The present invention describes an isolated antibody that binds to one of the human PRO proteins designated PRO212, PRO290, PRO341, PRO535, PRO619, PRO717, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009, PRO1025, PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184, PRO1187, PRO1281, PRO1293, PRO1309, PRO1317, PRO1710, PRO2094, PRO2145 OR PRO2198. PRO antagonists can be used to inhibit tumour cell growth. The PRO polypeptides and nucleotides are useful in the treatment, diagnosis and prevention of cancer. The antibodies and other anti-tumour compounds may be used to treat various conditions, including those characterised by overexpression and/or activation of the amplified PRO genes. Exemplary conditions or disorders to be treated with such antibodies and other compounds include benign or malignant tumours (e.g., renal, liver, kidney, bladder, breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulva, thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and neck tumours), leukaemias and lymphoid malignancies, other disorders such as neuronal, glial, astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal and blastocoelec disorders, and inflammatory, angiogenic and immunologic disorders. AAC58242 to AAC59366 represent PCR primers and hybridisation probes used in the isolation of the human PRO sequences. AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human PRO polynucleotide and protein sequences given in the exemplification of the present invention

SQ Sequence 1509 BP; 297 A; 451 C; 429 G; 332 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-46e-124 Length: 1509
 Score: 1258.50 Matches: 250
 Percent Similarity: 63.23% Conservative: 75
 Best Local Similarity: 48.64% Mismatches: 138
 Query Match: 47.67% Indels: 51
 DB: 3 Gaps: 7

US-09-303-232-6 (1-501) x AAC58395 (1-1509)

QY 8 LeuAlaLeuLeuAla---LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26

Db 28 CTGGCGCTGGCGCGCTCGCTCTCTGCGAGTGTCTGCAAGCGGAGTTCAGAGAGAGCTT 87

QY 27 LeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46

Db 88 TACAAGAGAGCTGGTCAAGAACTACAATCCCTTGGAGAGGCCCGTGGCAATGACTCGCAA 147

QY 47 ProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleIleAspValAspGluLysAsn 66

Db 148 CCACTCACCGTCTACTTCTCCCTGAGCCCTCTGCAGATCATGCGTGGATGAGAGAAC 207

QY 67 GlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTrp 86

Db 208 CAAGTTTAAACCAACCAACATTTGGCTGCAAAATGCTTGGACAGATCACTATTACAGTGG 267

QY 87 AsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsnLysLeuTrpLys 106

Db 268 AATGTGTCAGAAATATCCAGGGGTGAAGACTGTTCGTTTCCAGATCGCCAGATTTCGAAA 327

QY 107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126

Db 328 CCAGACATTTCTCTATACAGTGTGTGAGAGCGCTTTGACGCCACATCTCCACATTAAC 387

QY 127 ValValValArgSerGlyGlySerCysLeuTyrValProGlyIlePheLysSerThr 146

Db 388 GTGTTGGTGAATTTCTTGGGCAATTCGCCAGTACCTGCCCTCCAGGCATATTCAAGAGTTC 447

QY 147 CysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMetLysPheGly 166

Db 448 TGTACATCGATGTACGCTGGTTCCTCTTGATGTGCAGACTGCAAACTGAAGTTTGG 507

QY 167 SerThrThrTyrAspGlyAsnGlnLeuValLeuValLeuLysAspGluAlaGlyGlyAsp 186

Db 508 TCCTGTCTTACGGAGGCTGGTCTCTTGGATCTGCAGATCGAGAG-----GCAGAT 558

```
QY 187 LeuSerAspPheIleThrAenGlyGluTrpIleuLeuGlyMetProGlyLysLysAsn 206
Db 559 ATCAGTGGCTATATCCCAATGGAGATGGACCTAGTGGGAATCCCGGCAAGAGGAGT 618
QY 207 ThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMetIle 226
Db 619 GAAAGGTTCTATGAGTCTGCAAGAGCCCTACCCCGATGTCACCTTCACAGTGACCATG 678
QY 227 ArgArgArgThrLeuTyrPhePheAsnLeuIleValProCysValLeuIleSerSer 246
Db 679 CGCCGAGGAGCTCTACTATGGCTCAACCTGTGATCCCTGTGTGCTCATCTCCGCC 738
QY 247 MetAlaLeuLeuGlyPheThrLeuProAspSerGlyGluLysLeuThrLeuGlyVal 266
Db 739 CTGCGCTGTGGTTCCTCTCTCTGAGATTCGGGGAGAGATTTCCTCGGGGATA 798
QY 267 ThrIleLeuLeuSerLeuThrValPheLeuAsnValAlaGluThrLeuProGlnVal 286
Db 799 ACAGTCTACTCTCTTACCGTCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 858
QY 287 SerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSer 306
Db 859 TCCGATTCCGTTACCTTATGAGCCAGTACTTCCAGACCATGATCATCTGCGGCTC 918
QY 307 SerValValLeuThrValValValLeuAsnTyrHisArgThrAlaAspIleHisGlu 326
Db 919 TCGTGTGTGTGAGCGTGATGCTGCTGAGTACCACACGACGACCGCGGGGCAAG 978
QY 327 MetProGlnTrpIleLysSerValPheLeuGlnTrpLeuProTyrIleLeuArgMetSer 346
Db 979 ATGCCCAAGTGGACGAGTATCTCTTGAACCTGGTGGCGTGTCTCTCGAATGAAG 1038
QY 347 ArgProGlyLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu 366
Db 1039 AGGCCCGGGGAGGACAGAGTGGCGCGCGCTGCCAGCACAGCAGCGGCTGCGAGCTG 1098
QY 367 GluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAsp 386
Db 1099 GCCAGTGTGGAGATGAGCGCGTGGCG-----1125
QY 387 PheArgHisGlyProProProAsnSerThrAlaSerThrGlyAsnLeu-----403
Db 1126 -----CCGCGCGCC-----GCCAGCAACGGGAACCTCTGTATCATC 1161
QY 404 -----GlyProGlyCys 407
Db 1162 GGCTTCGCGCGCTGGAGCGGCTGCATGTCTCCGACCCCGACTCTGGGGTAGTGTGT 1221
QY 408 SerIlePheArgThrAspPheArgSerPheValArgProSerThrMetGluAspVal 427
Db 1222 GGC-----CGCATGGCTGTCTCCCGACGCGACGATGAGCAGCTC 1260
QY 428 -----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeuIleLeuArgGluLeu 445
Db 1261 CTGCACGGTGGGCAACCGCGAGGGGACCGGACTTGGCCAGATCTCTGGAGAGGTC 1320
QY 446 GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuIleSerAspTrp 465
Db 1321 CGCTACATTGCAACCGCTTCGGTCCGTCGAGACGAGGAGCGGGTCTGCGAGCGAGTGG 1380
QY 466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485
Db 1381 AAGTTCGCGCGCTGTGTGGTGGACCGCGCTGTGCTCATGGCTTCTCGGTCTTCCATC 1440
QY 486 IleAlaThrValAlaValLeuSerAlaProHisIleIle 499
Db 1441 ATCTGCATCCGCGATCTCTGATGTGGCTCCCAACTTCGTG 1482
RESULT 7
ID AAC90380
XX AAC90380 standard; cDNA; 1509 BP.
AC AAC90380;
```

```
XX 14-MAR-2001 (first entry)
DT Wild-type human alpha7 ligand gated ion channel coding sequence.
DE Human; alpha7 nicotinic acetylcholine gated ion channel;
KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance; ss.
XX Homo sapiens.
XX W0200073431-A2.
XX 07-DEC-2000.
XX 25-MAY-2000; 2000WO-US011862.
XX 27-MAY-1999; 99US-0136174P.
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX Groppi VE, Wolfe ML, Berkenpas MB;
XX WPI; 2001-061524/07.
XX P-PSDB; AAB50012.
XX Special cell culture medium for treating cells and for inducing mammalian
XX cell lines to conduct calcium ions, comprising specified concentrations
XX of ions of sodium, calcium and potassium at specified pH.
XX Example 5; Page 60-61; 77pp; English.
XX The present sequence is the coding sequence for wild-type human alpha7
XX nicotinic acetylcholine gated ion channel. The human alpha7 ion channel
XX was used in the construction of an alpha7/5-hydroxytryptamine (5-HT3)
XX chimeric ligand gated ion channel (see AAC90382 and AAB50014). The
XX alpha7/5-HT3 chimeric ion channel can be expressed by recombinant cells
XX in the present invention, resulting in preferential calcium ion
XX conductance by the cells
XX Sequence 1509 BP; 297 A; 451 C; 429 G; 332 T; 0 U; 0 Other;

Alignment Scores:
Pred. NO.: 1.46e-124 Length: 1509
Score: 1258.50 Matches: 250
Percent Similarity: 63.23% Conservative: 75
Best Local Similarity: 48.64% Mismatches: 138
Query Match: 47.67% Indels: 51
DB: 4 Gaps: 7

US-09-303-232-6 (1-501) x AAC90380 (1-1509)
QY 8 LeuAlaLeuLeuAla---LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26
Db 28 CTGGCGTGGCGGTGGTCTCTCTGACGTGTCTCTGACGCGAGTCTCCAGCGAGTTC 87
QY 27 LeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46
Db 88 TACAAGGAGCTGGTCAAGAACTACATCCCTTGGAGAGCGCGTGGCCATGACTCGCAA 147
QY 47 ProLeuGluValArgPheGlyLeuThrLeuGlnIleIleAspValAspGluLysAsn 66
Db 148 CCACCTACCGTCTCTCTCTCTGAGCTCTCTGAGATCATGTCGTCGTCGTCGTCGTCG 207
QY 67 GlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTrp 86
Db 208 CAAGTTTACCAACACATTTGGTGCATATGTCTTGGACAGATCACTATTACAGTGG 267
QY 87 AsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsnLysLeuTrpLys 106
Db 268 AATGTGTCAAGATATCCAGGGGTGAAGACTCTTCTGTTTCCAGATGTCGTCGTCGTCG 327
QY 107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126
Db 107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126
```

328 CCAGACATCTCTCTATACAGTGTCTGATGAGCGCTTTGACGCCACATTCACACATAAC 387
127 ValValValArgSerGlyGlySerCysLeuTyrValProProGlyIlePhePheSerThr 146
388 GTGTGTGTGAATCTCTGGGCATTTGCCAGTACCTGCTCCAGGCATATTCAGAGTTCC 447
147 CysLysMetAspIleAlaTyrPheProPheAspGlnHisCysAspMetLeuPheGly 166
448 TGCTACATGATGTACGCTGTGTTCCTTTGATGTGCAGCACTGCAAACTGAAGTTGG 507
167 SerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyGlyAsp 186
508 TCCTGGTCTTACGGAGGCTGTCTCTGATCTGCAGATGCAGAG-----GCAGAT 558
187 LeuSerAspPheIleThrAsnGlyGluTyrPheLeuIleGlyMetProGlyLysAsn 206
559 ATCAGTGGCTATATCCCAATGGAGATGGACCTAGTGGAAATCCCGGCAAGAGGAT 618
207 ThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMetIle 226
619 GAAGGTTCTATGATGCTGCAAGAGCCCTACCCGATGTCACTTCACAGTGCACCATG 678
227 ArgArgArgThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleSerSer 246
679 CGCGCAGGACGCTCTACTATGCGCTCAACCTGCTGATCCCTGTGTCTCATCTCCGCC 738
247 MetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThrLeuGlyVal 266
739 CTGCGCTCTGTGTGTCTCTGCTTCTCGCATTCCTGGGAGATTCCTGGGAGATA 798
267 ThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnVal 286
799 ACAGTCTTACTCTCTTACCGTCTTATGCTGTCTGCTGTGCTGATCATGCTGGGCTC 858
287 SerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSer 306
859 TCCGATTTCGTATCATTCATGACCCAGTACTTCCGACGACCATGATCATCTGGGGCTC 918
307 SerValValLeuThrValValValLeuAsnTyrHisArgThrAlaAspIleHisGlu 326
919 TCGTGTGTGTGAGGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978
327 MetProGlnTrpIleLysSerValPheLeuGlnTrpLeuProTrpIleLeuArgMetSer 346
979 ATGCCAAGTGGACGACGATCTCTCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1038
347 ArgProGlyLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu 365
1039 AGGCCCGGGAGGACCAAGGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1098
367 GluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAsp 386
1099 GCCAGTGTGAGATGACGCGCGTGGCG-----1125
387 PheArgHisGlyProProProAsnSerThrAlaSerThrGlyAsnLeu-----403
1126 -----CGCGCGCC-----GCCAGCAACGGGAACCTGCTGTATCATC 1161
404 -----GlyProGlyCys 407
1162 GGCTTCGCGCGCTGGACGCGTGCATGTGCTCCGACCCCGCTGCTGGGTAGTGTGT 1221
408 SerIlePheArgThrAspPheArgArgSerPheValArgProSerThrMetGluAspVal 427
1222 GGC-----CGCATGCGCTGCTCCCGCCACGACGATGACGACCTC 1260
428 -----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeuIleLeuArgGluLeu 445
1261 CTGACGCGGGCAACCCCGAGGGGACCGGATGTCGCAAGATCTCGGAGGAGGTC 1320
446 GlnPheIleThrAlaArgMetLysLysAlaAspGluAlaGluLeuIleSerAspTrp 465
1321 CGTACATTCCTCAATCGCTTCCGCTGCTGACGAGCAAGAGCGGCTGCTGACGAGTGG 1380

Qy 466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485
Db 1381 AAGTTCCGCGCTGTGTGTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Qy 486 IleAlaThrValAlaValLeuLeuSerAlaProHisIleIle 499
Db 1441 ATCTGCACCATCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1482

RESULT 8

AAV12197
ID AAV12197 standard; cDNA; 1876 BP.
XX AAV12197;
XX 14-MAY-1998 (first entry)
XX Human neuronal nicotinic acetylcholine receptor alpha-7 subunit cDNA.
XX Human; neuronal nicotinic acetylcholine receptor; alpha-7 subunit;
KW brain tissue; screening; NACHR; antibody; ds.
XX Homo sapiens.

Key Location/Qualifiers
CDS 73..1581
FT /*tag= a
FT /product= "neuronal nicotinic acetylcholine receptor
FT alpha-7 subunit"

XX WO9420617-A2.
PN 15-SEP-1994.
XX 08-MAR-1994; 94WO-US002447.
XX 08-MAR-1993; 93US-00028031.
XX (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
XX Elliott KJ, Ellis SB, Harpold MM;
XX WPI; 1994-303024/37.
XX P-PSDB; AAM44153.

XX Human neuronal nicotinic acetylcholine receptor subunits and DNA - also transformed cells useful for screening cpds. which modulate activity of the receptor.

XX Claim 8; Page 78-79; 99pp; English.

XX The present sequence encodes a human neuronal nicotinic acetylcholine receptor (NACHR) subunit. The cells expressing the alpha and/or beta NACHR subunits may be used in a method of screening compounds to identify any which modulate the activity of human neuronal NACHR. Subunit specific antibodies may be used to monitor the distribution and expression density of various subunits in normal vs diseased brain tissues. Testing of single receptor subunits or specific receptor subunit combinations with a variety of potential agonists or antagonists provides information with respect to the function and activity of the individual subunits and should lead to the identification and design of compounds that are capable of very specific interaction with one or more receptor subtypes. The resulting drugs should exhibit fewer unwanted side effects than drugs identified e.g. screening with cells that express a variety of subtypes

XX Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 2,02e-124 Length: 1876
Score: 1258.50 Matches: 250
Percent Similarity: 63.23% Conservative: 75
Best Local Similarity: 48.64% Mismatches: 138
Query Match: 47.67% Indels: 51

```

DB:      2          Gaps:      7
US-09-303-232-6 (1-501) x AAV12197 (1-1876)

Qy      8 LeuAlaLeuLeuAla----LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26
      100 CTGGCGTGGCGCGTGCCTCTCGAGTGTCCCTGCAAGCGAGTTCACAGAGGAAGCTT 159
      27 LeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46
      160 TACAAGAGAGCTGGTCAAGAACTACAATCCCTTGGAGAGGCCGTGGCCATGACTCCGAA 219
      47 ProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleLeuAspValAspGluLysAsn 66
      220 CCATCTACCGTCTACTTCTCCCTGAGCTCTCTGAGATCATGGAGCTGGATGAGAAGAC 279
      67 GlnLeuLeuLeuThrAsnIleThrLeuSerLeuThrAsnGluThrProAsnLysLeuArgTyr 86
      280 CAAGTGTTCACCAACCAATTTGGCTGCAATGTCTTGGACAGATCATCTATTACAGTGG 339
      87 AsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsnLysLeuThrLys 106
      340 AATGTGTGAGAATATCCAGGGGTGAAGACTTCTGTTTCCAGATGGCCAGATTGGAAA 399
      107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126
      400 CCAGACATCTCTCTATAACAGTGTCTGATGAGCGCTTTGACGCCACATCCACACTAAC 459
      127 ValValValArgSerGlyGlySerCysLeuTyrValProGlyIlePheLysSerThr 146
      460 GTGTTGGTGAATTTCTTGGGCAATTCAGTACCTGCTCCAGGCAATATTCAGAGTTTCC 519
      147 CysLysMetAspIleAlaThrPheProPheAspAspGlnHisCysAspMetLysPheGly 166
      520 TGCTACATCGATGACGTGTGTTTCCCTTTGATGTCAGCACTGCAAACTGAAGTTTGGG 579
      167 SerTyrThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyLys 186
      580 TCCTGGTCTTACGGAGGCTGCTTGGATCTGCAGATGCAGGAG-----GCAGAT 630
      187 LeuSerAspPheIleThrAsnGlyGluThrTyrLeuIleGlyMetProGlyLysLysAsn 206
      631 ATCAGTGGCTATATCCCAATGGAGATGGGACCTAGTGGGAATCCCGGCAAGAGAGT 690
      207 ThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMetIle 226
      691 GAAAGTTCTATGAGTGTCTGAAGAGCCCTACCCCGATGACCTTCACAGTGACCAATG 750
      227 ArgArgArgThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleSerSer 246
      751 CGCGCAGACGCTCTACTATGGCTCAACCTGTGATCCCTGTGTGCTCATCTCCGCC 810
      247 MetAlaLeuLeuGlyPheThrLeuProAspSerGlyGluLysLeuThrLeuGlyVal 266
      811 CTGCGCCTGTGTGTTCTCTGCTTCCCTGATTCGGGGAGAGATTTCCTGGGGATA 870
      267 ThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnVal 286
      871 ACAGTCTTACTCTCTTACCGTCTTCATGTGCTGCTGGCTGAGATCATGCCGCCAACA 930
      287 SerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSer 306
      931 TCCGATTCCGTACCACTTATGATCCCAAGTACTTCGGCAGCAACCATCATCATCTGGGCC 990
      307 SerValValLeuThrValValValLeuAsnTyrHisArgThrAlaAspIleHisGlu 326
      991 TCGGTGGTGGTACCGGTATCGTGTGAGTACCAACCAACAGCCAGCCCGGAGGGGCAAG 1050
      327 MetProGlnTyrIleLysSerValPheLeuGlnTyrLeuProTyrIleLeuArgMetSer 346
      1051 ATGCCCAAGTGACACAGATCATCTCTGAACTGGTGGCGGTCTCTCGAATGAAG 1110
      347 ArgProGlyLysLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu 366

```

RESULT 9

```

AAAT48239
ID   AAT48239 standard; DNA; 1876 BP.
XX
AC   AAT48239;
XX
DT   09-APR-1997 (first entry)
DE   Neuronal nicotinic acetylcholine receptor alpha-7 subunit DNA.
XX
KW   Neuronal nicotinic acetylcholine receptor; nAChR; neurotransmitter;
    ligand-gated receptor; ds.
XX
OS   Homo sapiens.
XX
FH   Key Location/Qualifiers
    CDS 73..1581
        /*tag= a
XX
XX   WO9641876-A1.
XX
PD   27-DEC-1996.
XX
PF   07-JUN-1996; 96WO-US009775.
XX
PR   07-JUN-1995; 95US-00484722.
XX
PA   (SIBI-) SIBIA NEUROSCIENCES INC.
XX
PI   Elliott KJ, Harpold MM;
XX
XX   WPI; 1997-065463/06.
DR   P-PSDB; AAW09025.
XX
XX   Nucleic acids encoding nicotinic acetylcholine receptor sub-units - used
    in screening to determine the effect of drugs on the receptor.
XX
PS   Disclosure; Page 71-73; 108pp; English.

```

XX A DNA sequence (AAT48239) codes for the alpha-7 subunit (AAW09025) of the
 CC human neuronal nicotinic acetylcholine receptor (nAChR). Host cells, esp.
 CC mammalian cells or amphibian oocytes, carrying alpha-7 nucleic acids,
 CC opt. in combination with other alpha and/or beta subunit nucleic acids
 CC (see also AAT48232-38, AAT48240-41), express recombinant nAChR subunits
 CC useful for identifying cpds. that modulate the activity of human nAChRs
 XX

SQ Sequence 1876 BP; 369 A; 553 G; 531 C; 423 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,02e-124 Length: 1876
 Score: 1258.50 Matches: 250
 Percent Similarity: 63.23% Conservative: 75
 Best Local Similarity: 48.64% Mismatches: 138
 Query Match: 47.67% Indels: 51
 DB: 2 Gaps: 7

US-09-303-232-6 (1-501) x AAT48239 (1-1876)

Qy 8 LeuAlaLeuLeuAla--LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26
 Db 100 CTGGCGCTGGCGCGTCTCTCTGACGTGCTCCGCAAGCGGAGTTCACGAGGAAGCTT 159
 Qy 27 LeuAsnAlaLeuLeuAlaAsnThrLeuGluArgProValAlaAsnGluSerGlu 46
 Db 160 TACAAGGAGCTGGTCAAGAAGTACAAATCCCTTGGAGAGCGCGTGGCCATGATCTGCAA 219
 Qy 47 ProLeuGluValArgPheGlyLeuThrLeuGlnGlnLeuLeuLeuLeuLeuLeu 66
 Db 220 CCACCTCCCGTCTACTTCTCCCTGAGCTCTGAGGAGCGCGTGGATGAGAGAAC 279
 Qy 67 GlnLeuLeuLeuThrAsnLeuTrpLeuSerLeuGluTrpAsnAspThrLeuArgTrp 86
 Db 280 CAAGTTTAAACCAACCAATTTGGCTGCAATGCTTGGACAGATCACTATTATGAGTGG 339
 Qy 87 AsnAspSerGluTrpGlyGlyValLysAspLeuArgGlnLeuThrProAsnLysLeuTrpLys 106
 Db 340 AATGTGTCAGAAATATCCAGGGGTGAGACTGTTGTTTCCAGATGGCCAGATTGGAAA 399
 Qy 107 ProAspValLeuMetThrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126
 Db 400 CCAGACATCTCTCTATACAGTGTGATGAGCGCTTTGACGCCACATTCACACATAAC 459
 Qy 127 ValValValArgSerGlyGlySerCysLeuTrpValProProGlyIlePheLysSerThr 146
 Db 460 GTGTGTGTGAATCTCTGGGCATTTGGCAGTACCTCCAGCATATTTCAAGAGTTCC 519
 Qy 147 CysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMetLysPheGly 166
 Db 520 TGCTACATCGATGTACGCTGGTTCCTTTGATGTGCAGACTGCAAACTGAAGTTTGGG 579
 Qy 167 SerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyLys 186
 Db 580 TCCTGTGCTTACGGAGCGTGGTCTGATCTGAGATGCGAGAG-----GCAGAT 630
 Qy 187 LeuSerAspPheIleThrAsnGlyGluTrpTyrLeuLeuGlyMetProGlyLysLysAsn 206
 Db 631 ATCAGTGGCTATATCCCAAGAGATGGAGCTAGTGGGAATCCCGGCAAGAGAGT 690
 Qy 207 ThrIleThrTyrAlaCysCysProGluProTrpValAspValThrPheThrIleMetIle 226
 Db 691 GAAAGGTTCTATGATGCTGCAAGAGCCCTACCCGATGTCCACTTCACAGTACCATG 750
 Qy 227 ArgArgArgThrLeuTyrTyrPhePheAsnLeuLeuValProCysValLeuLeuSerSer 246
 Db 751 CGCGCAGGAGCGCTCTACTATGGCTCAACCTGCTGATCCCTGTGTCTCATCTCCGCC 810
 Qy 247 MetAlaLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThrLeuGlyVal 266
 Db 811 CTCGCCCTGTGGTGTCTCTGCTGCTGCAATTCGGGAGAGATTCCTCCCTGGGATA 870
 Qy 267 ThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnVal 286

Db 871 ACAGTCTACTCTCTTACCGTCTTATGCTGTGTGGGTGAGATCATGCCGCCAACAA 930
 Qy 287 SerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSer 306
 Db 931 TCCGATTTCGTAACCATGATAGCCAGTACTTCGCCAGCACCATGATCATCTGTGGCCTC 990
 Qy 307 SerValValLeuThrValValValLeuAsnTyrHisArgThrAlaAspIleHisGlu 326
 Db 991 TCGGTGTGTGTGACGGTGATCGTGTGAGTACCACACACACACCGCGCGGCGCAAG 1050
 Qy 327 MetProGlnTrpIleLysSerValPheLeuGlnTrpLeuProTrpIleLeuArgMetSer 346
 Db 1051 ATGCCCAAGTGGACGAGATCATCTCTGAACTGGTGGCGGTCTCTCGCAATGAAG 1110
 Qy 347 ArgProGlyLysLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu 366
 Db 1111 AGGCCCGGGAGGACAAAGGTGCGCCGCGCTGCCAGACAGCAGCGCGGTGCAGCCTG 1170
 Qy 367 GluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAsp 386
 Db 1171 GCCAGTGTGGAGATGAGCGCGCGTGGCG-----1197
 Qy 387 PheArgHisGlyProProProAsnSerThrAlaSerThrGlyAsnLeu-----403
 Db 1198 -----CCGCCGCC-----GCCAGCAACGGGAACCTCTGTATCATC 1233
 Qy 404 -----GlyProGlyCys 407
 Db 1234 GGCTTCGCGCGCTGGAGCGGTGCACCTGTGTCCGACCCCGACTCTGGGGTAGTG 1293
 Qy 408 SerIlePheArgThrAspPheArgSerPheValArgProSerThrMetGluAspVal 427
 Db 1294 GGC-----CGCATGCGCTGTCTCCCGACGACGATGAGCACCTC 1332
 Qy 428 -----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeuLeuLeuArgGluLeu 445
 Db 1333 CTGCACGCGCGGCAACCCCGAGGGGACCGGACTTGGCCAGATCTCTGGAGAGGTC 1392
 Qy 446 GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuLeuSerAspTrp 465
 Db 1393 CGCTACATTGCCAATCGTTCCTCCGTCGACGAGCAAGAGCGGGCTCTGCAGCGAGTGG 1452
 Qy 466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485
 Db 1453 AAGTTCGCCCGCTGTGTGTGGACCGCTGTGCTCATGGCTTCTCGTCTTCCCATC 1512
 Qy 486 IleAlaThrValAlaValLeuLeuSerAlaProHisIleIle 499
 Db 1513 ATCTGCACCATCGCATCTTGTGTGGCTCCCAACTTCGTG 1554

RESULT 10
 ABS54875
 ID ABS54875 standard; cDNA; 1876 BP.
 AC ABS54875;
 XX
 XX 06-DEC-2002 (first entry)
 DT
 XX Human neuronal nicotinic acetylcholine receptor alpha 7 subunit cDNA.
 DE
 XX Human; neuronal nicotinic acetylcholine receptor; nNACHR; gene; ss;
 KW ion flux; alpha 7 subunit.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 PH 73..1581
 FT /*tag= a
 FT /product= "Human nNACHR alpha 7 subunit"
 XX
 PN US6440681-B1.


```

Db 1171 GCCAGTGTGGAGATGAGCGCGTGGCG----- 1197
Qy 387 PheArgHisGlyProProProAsnSerThrAlaSerThrGlyAsnLeu----- 403
Db 1198 -----CCGCGCGCC-----GCCAGCAACGGAACCTCTGTATCATC 1233
Qy 404 -----GlyProGlyCys 407
Db 1234 GGCTTCGCGCGCTGGAGCGGTGCACCTGTCTCCGACCCCGACTCTGGGGTAGTGTGT 1293
Qy 408 SerIlePheArgThrAspPheArgSerPheValArgProSerThrMetGluAspVal 427
Db 1294 GGC-----CGCATGGCTGTCTCCCGACGCGACGATGAGCACCTC 1332
Qy 428 -----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeuLeuArgGluLeu 445
Db 1333 CTGCACGCGCGGCAACCCCGAGGGGACCCCGGACTTGGCCAAAGATCCTCGAGGAGTCT 1392
Qy 446 GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuIleSerAspTrp 465
Db 1393 CGCTACATTGCCAATCGCTTCGCTCCGACGAGCAAGAGCGAGCGGTCTGCAGCGAGTGG 1452
Qy 466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485
Db 1453 AAGTTCGCGCGCTGTGTGTGGACCGCGCTGTGCCTCATGGCGCTTCTCGGTCTTCCACCATC 1512
Qy 486 IleAlaThrValAlaValLeuLeuSerAlaProHisIleIle 499
Db 1513 ATTCGACCATCGGCATCCTGATGCGGCTCCCAACTTCGTG 1554

RESULT 12
ID ADA10864
XX ADA10864 standard; DNA; 1876 BP.
AC ADA10864;
XX
XX 06-NOV-2003 (first entry)
XX
XX Human neuronal nicotinic acetylcholine receptor alpha 7 subunit DNA.
KW ds; gene; alpha 7 subunit; human;
KW neuronal nicotinic acetylcholine receptor; ligand-gated ion channel;
KW synaptic transmission; gene therapy; transgenic.
XX
XX Homo sapiens.
FH Key Location/Qualifiers
FT 5'UTR 1..72
FT FT /*tag= a
FT CDS 73..1584
FT FT /*tag= b
FT FT /product= "nAChR alpha 7 subunit"
FT 3'UTR 1585..1876
FT FT /*tag= c

US6524789-B1.
XX
XX
XX 25-FEB-2003.
XX
XX 07-JUN-1996; 96US-00660451.
XX
XX 07-JUN-1995; 95US-00484722.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Elliott KJ, Harpold MM;
XX
XX WPI; 2003-511917/48.
XX
XX P-PSDB; ADA10874.
XX
XX New nucleic acid encoding an alpha-6 or a beta-3 subunit of a human
XX neuronal nicotinic acetylcholine receptor (nAChR), useful for identifying
XX compounds that modulate human neuronal nAChR activity.

```

```

XX Claim 33; Col 67-72; 63pp; English.
PS
XX
XX The invention relates to an isolated nucleic acid molecule comprising a
CC nucleotide sequence encoding an alpha 6 or a beta 3 subunit of a human
CC neuronal nicotinic acetylcholine receptor nAChR. nAChR's form ligand-
CC gated ion channels that mediate synaptic transmissions between nerve and
CC muscle and between neurons upon interaction with the neurotransmitter
CC acetylcholine. The nucleic acid molecule is useful for identifying
CC compounds that modulate human neuronal nAChR. The present sequence
CC represents DNA encoding the human neuronal nicotinic acetylcholine
CC receptor, nAChR, alpha 7 subunit. Note: the present sequence decodes to
CC the protein shown in ADA10874 not the one described in the specification
CC as being the nAChR alpha 7 subunit ADA10865.
XX
SQ Sequence 1876 BP; 369 A; 553 C; 531 G; 423 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,02e-124 Length: 1876
Score: 1258.50 Matches: 250
Percent Similarity: 63.23% Conservative: 75
Best Local Similarity: 48.64% Mismatches: 138
Query Match: 47.67% Indels: 51
DB: 8 Gaps: 7

US-09-303-232-6 (1-501) x ADA10864 (1-1876)
Qy 8 LeuAlaLeuLeuAla---LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26
Db 100 CTGGCGCTGGCGCGCTCGCTCCTCGACGTGTCCCTGCAAGCGAGTTCACAGGAAGCTT 159
Qy 27 LeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46
Db 160 TACAAGGAGCTGGTCAAGAACTACAATCCCTTGGAGAGGCCGCTGGCAATGACTCGCA 219
Qy 47 ProLeuGluValArgPheGlyLeuThrLeuGlnIleIleAspValAspGluLysAsn 66
Db 220 CCATCTACCGCTCTACTTCTCCCTGAGCTCTCTGACATCATGGACGTGATGAGAAGAC 279
Qy 67 GlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTrp 86
Db 280 CAAGTTTAAACCAACATTTGGCTGCAATGTCTTGGACAGATCTATTTACGATGG 339
Qy 87 AsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsnLysLeuTrpLys 106
Db 340 AATGTGTCAAGATATCCAGGGGTGAAGACTGTCTTCCAGATGGCCAGATTTGGAAA 399
Qy 107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126
Db 400 CCAGACATTTCTTCTATACAGTGTCTGATGAGCGCTTTGACGCCACATTCACACTAAC 459
Qy 127 ValValValArgSerGlyGlySerCysLeuTyrValProGlyIlePheLysSerThr 146
Db 460 GTGTGTGGTAATCTTCTGGGCATTCGACGATTCCTCCCTCAGGATATTCAGAGTTCC 519
Qy 147 CysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAspMetLysPheGly 166
Db 520 TGCTACATCGATGATACGCTGGTTTCCCTTTCATGTGACGACTGCAAACTGAAGTTGGG 579
Qy 167 SerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyClyAsp 186
Db 580 TCCTGGTCTTACGGAGGCTGGTCTCTGGATCTGCAGATGCAGGAG-----GCAGAT 630
Qy 187 LeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGlyMetProGlyLysLysAsn 206
Db 631 ATCAGTGGCTATATCCCATGGAGATGGACCTAGTGGGAATCCCGCGGAAGAGGAGT 690
Qy 207 ThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMetIle 226
Db 691 GAAAGGTTCTATGATGCTGCAAGAGGCCCTTACCCGATGTCACCTTCACAGATGACCATG 750
Qy 227 ArgArgArgThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleSer 246

```

```

Db 751 CGCGCAGGAGCGCTCTACTATGGCTCAACCTGCTGATCCCTGTGTGCTCATCTCCGCC 810
Qy 247 MetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuThrLeuGlyVal 266
Db 811 CTGCCCTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 870
Qy 267 ThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnVal 286
Db 871 ACAGTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 930
Qy 287 SerAspAlaIleProLeuLeuGlyThrTyPheAsnCysIleMetPheMetValAlaSer 306
Db 931 TCCGATTCCGTTACCATGTATGACCCAGTACTTCCAGCACCACCATGATCATCTGTGGGCTC 990
Qy 307 SerValValLeuThrValValValLeuAsnTyRhiShiArgThrAlaAspIleHisGlu 326
Db 991 TCGGTGTGTGTGTCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1050
Qy 327 MetProGlnTrpIleLysSerValPheLeuGlnTrpLeuProTrpIleLeuArgMetSer 346
Db 1051 ATGCCCAAGTGGACCAAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1110
Qy 347 ArgProGlyLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu 366
Db 1111 AGGCCCGGGAGGACAAAGTGGCGCGCGCTGCCAGCACAGCAGCGCGCTGCAGCCTG 1170
Qy 367 GluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAsp 386
Db 1171 GCCAGTGTGAGATGAGCGCGCTGGCG-----GlyProGlyCys 407
Qy 387 PheArgHisGlyProProProAsnSerThrAlaSerThrGlyAsnLeu----- 403
Db 1198 -----CGCGCGCC-----GCCAGCACGGGAACTCTGCTGATCATC 1233
Qy 404 -----GlyProGlyCys 407
Db 1234 GGCTTCCGCGCGCTGGACGGGTGCTGTGTCCGACCCCGACTCTGGGGTAGTGTGT 1293
Qy 408 SerIlePheArgThrAspPheArgSerPheValArgProSerThrMetGluAspVal 427
Db 1294 GGC-----CGCATGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1332
Qy 428 -----GlyGlyGlyLeuGlySerHisArgGluLeuHisLeuLeuArgGluLeu 445
Db 1333 CTGCACGGCGGGNACCCCGCGGGGACCCGACTTGGCCAGATCTCTGGAGAGTCTC 1392
Qy 446 GlnPheIleThrAlaArgMetLysAlaAspGluAlaGluLeuLeuSerAspTrp 465
Db 1393 CGTACATTGCCAATCGCTTCCGCTGCCAGCAGGAGCGGGTCTGCAGCGAGTGG 1452
Qy 466 LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle 485
Db 1453 AAGTTCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1512
Qy 486 IleAlaThrValAlaValLeuSerAlaProHisIleIle 499
Db 1513 ATCTGCACCATCGCATCTCTGATGTGGCTCCCACTCTGTG 1554

```

RESULT 13

AAV44687

ID AAV44687 standard; cDNA; 1590 BP.

AC AAV44687;

XX 09-OCT-1998 (first entry)

XX V274T variant human alpha7 nAChR coding sequence.

XX Alpha7 nAChR; alpha7 nicotinic acetylcholine receptor subunit; cancer;
 KW neurodegeneration; enzyme dysfunction; affective disorder; therapy;
 KW immune dysfunction; diabetic neuropathy; Alzheimer's disease;
 KW schizophrenia; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 9..1517

XX FT /*tag= a

XX WO9828331-A2.

XX PN 02-JUL-1998.

XX PD 22-DEC-1997; 97WO-US023405.

XX PF 20-DEC-1996; 96US-00771737.

XX PR (ABBO) ABBOTT LAB.

XX PA Briggs CA, Gopalakrishnan M, McKenna DG, Monteggia LM, Roch J;

XX PI Sullivan JP, Touma E;

XX DR WPI; 1998-377593/32.

XX DR P-PSDB; AAW69216.

XX PT Nucleic acid encoding variant of human alpha 7 nicotinic acetylcholine

XX PT receptor sub-unit - used to identify modulators of the receptor,

XX PT potentially useful for treating neuro-degeneration, cancer etc.

XX PS Claim 14; Fig 2; 44pp; English.

XX CC This sequence encodes the V247T variant of human alpha7 nicotinic

XX CC acetylcholine receptor (nAChR) subunit of the invention. Cells containing

XX CC the DNA are used to express the protein and to identify modulators of

XX CC alpha7 nAChR activity or cytoprotective agents, e.g. antisense compounds

XX CC or antagonists that are potentially useful for treating

XX CC neurodegeneration, enzyme dysfunction, affective disorders and immune

XX CC dysfunction, such as cancer, post-herpetic neuralgia, diabetic

XX CC neuropathy, osteoarthritis, Alzheimer's or Parkinson's diseases, kuru,

XX CC psychosis and schizophrenia. Probes based on the DNA are used to detect

XX CC the DNA in usual hybridisation or amplification tests, while monoclonal

XX CC antibodies are used to detect the protein for diagnosis (in vitro or by

XX CC in situ immuno-fluorescent assay). Compared with wild-type alpha7 nAChR,

XX CC the protein has about 100-fold greater sensitivity to cholinergic

XX CC receptor agonists (nicotine or acetylcholine) and response to these

XX CC agonists decays more slowly, but the wild-type inward rectification is

XX CC retained

XX SQ Sequence 1590 BP; 322 A; 471 C; 449 G; 348 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4,24e-124	Length:	1590
Score:	1254.50	Matches:	249
Percent Similarity:	63.04%	Conservative:	75
Best Local Similarity:	48.44%	Mismatches:	139
Query Match:	47.52%	Indels:	51
DB:	2	Gaps:	7

US-09-303-232-6 (1-501) x AAV44687 (1-1590)

Qy 8 LeuAlaLeuLeuAla---LeuLeuProValSerGluGlnGlyProHisGluLysArgLeu 26

Db 36 CTGGCACTGGCAGCATCT 95

Qy 27 LeuAsnAlaLeuLeuAlaAsnTyRAsnThrLeuGluArgProValAlaAsnGluSerGlu 46

Db 96 TACAAGAGCTGGTCTCAAGAACTCAATCCCTCGAGAGGCCCGTGGCAATGACTCGCAA 155

Qy 47 ProLeuGluValArgPheGlyLeuThrLeuGlnIleIleAspValAspGluLysAsn 66

Db 156 CCACTCACCGTCTACTTCT 215

Qy 67 GlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAspTyRAsnLeuArgTrp 86

Db 216 CAAGTTTTRACCCACCAACATTTGGCTGCTTGGACAGATCACTACTTATACAGTGG 275

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:31:03 ; Search time 12.1919 Seconds
(without alignments)
3952.801 Million cell updates/sec

Title: US-09-303-232-6
Perfect score: 2640
Sequence: 1 MAPMLAALALLPVSSEQ.....LFTIATVALLSAPHIVQ 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: Piri.*
2: Piri2.*
3: Piri3.*
4: Piri4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1260	47.7	502	1 ACHUA7	nicotinic acetylch
2	1255.5	47.6	502	2 A57175	nicotinic acetylch
3	1253	47.5	502	2 JN0113	nicotinic acetylch
4	1252.5	47.4	502	2 G02259	alpha 7 neuronal n
5	1246.5	47.2	502	2 T01378	nicotinic receptor
6	1237.5	46.9	498	2 S68588	nicotinic acetylch
7	1226.5	46.5	511	2 JH0173	alpha-bungarotoxin
8	1132	42.9	461	2 T25671	hypothetical prote
9	1083.5	41.0	542	2 T19862	hypothetical prote
10	1081	40.9	560	2 T19622	hypothetical prote
11	989	37.5	503	2 A53956	nicotinic acetylch
12	982.5	37.2	502	2 A37040	nicotinic acetylch
13	970.5	36.8	495	2 S60589	acetylcholine rece
14	970.5	36.8	499	2 A24572	nicotinic acetylch
15	964.5	36.5	567	1 ACFPA1	nicotinic acetylch
16	944	35.8	557	2 S12359	nicotinic acetylch
17	943	35.7	576	1 ACFPA2	nicotinic acetylch
18	929.5	35.2	494	2 T09289	nicotinic acetylch
19	924.5	35.0	528	1 ACCH2N	nicotinic acetylch
20	919.5	34.8	512	2 B37014	nicotinic acetylch
21	919	34.8	511	2 A40110	nicotinic acetylch
22	910	34.5	500	2 S12899	nicotinic acetylch
23	909.5	34.5	495	2 B35721	nicotinic acetylch
24	906.5	34.3	517	2 A30992	probable nicotinic
25	906	34.3	457	1 ACBOA1	nicotinic acetylch
26	905	34.3	457	1 ACCHU1	nicotinic acetylch
27	901	34.1	521	1 ACFENN	nicotinic acetylch
28	900	34.1	503	2 JH0174	nicotinic acetylch
29	897.5	34.0	627	2 JC4021	nicotinic acetylch

ALIGNMENTS

RESULT 1
ACHUA7

nicotinic acetylcholine receptor alpha-7 chain precursor, neuronal - human
N;Alternate names: cholinergic nicotinate receptor alpha-7 chain
C;Species: Homo sapiens (man)
C;Date: 12-Aug-1996 #sequence revision 31-Jan-1997 #text_change 22-Jun-1999
C;Accession: I37185; A54194; S60309
R;Peng, X.; Katz, M.; Gerzanich, V.; Anand, R.; Lindstrom, J.
Mol. Pharmacol. 45, 546-554, 1994
A;Title: Human alpha 7 acetylcholine receptor: cloning of the alpha 7 subunit from the s
omers expressed in Xenopus oocytes.
A;Reference number: I37185; MUID:94195283; PMID:8145738
A;Accession: I37185
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-502 <PEN>
A;Cross-references: EMBL:X70297; NID:9496606; PIDN:CAA49778.1; PID:9496607
A;Experimental source: brain neuroblastoma cell line SHSY-5Y
R;Chini, B.; Raimond, E.; Elgoyhen, A.B.; Moralli, D.; Balzaretti, M.; Heinemann, S.
Genomics 19, 379-381, 1994
A;Title: Molecular cloning and chromosomal localization of the human alpha-7-nicotinic
A;Reference number: A54194; MUID:94245214; PMID:8188270
A;Accession: A54194
A;Molecule type: mRNA

30	896	33.9	456	1	ACCHAN	nicotinic acetylch
31	895	33.9	502	2	S10505	nicotinic acetylch
32	891	33.8	622	1	ACCH4N	nicotinic acetylch
33	890	33.7	457	2	S13872	nicotinic acetylch
34	890	33.7	470	2	A39218	nicotinic acetylch
35	888	33.6	625	2	A26456	nicotinic acetylch
36	887	33.6	498	2	G02421	nicotinic acetylch
37	885.5	33.5	459	2	A14703	nicotinic acetylch
38	885	33.5	457	2	A24383	nicotinic acetylch
39	883.5	33.5	445	2	I49458	nicotinic acetylch
40	881	33.4	491	1	ACCHNN	nicotinic acetylch
41	879.5	33.3	461	2	I50548	acetylcholine rece
42	876.5	33.2	461	1	ACRFA1	nicotinic acetylch
43	875.5	33.2	457	2	A28529	nicotinic acetylch
44	869.5	32.9	457	2	S08162	nicotinic acetylch
45	862	32.7	559	2	E89134	protein F2566.4 [i

Query Match 47.7%; Score 1260; DB 1; Length 502;
Best Local Similarity 48.5%; Pred. No. 2.9e-99;
Matches 247; Conservative 75; Mismatches 137; Indels 50; Gaps 6;

A;Residues: 24-25,'ET',28-41,'X',43-45,'X',47 <CON>
C;Comment: This acetylcholine receptor is blocked by alpha-bungarotoxin and is localized
C;Genetics:
A;Introns: 19/1; 65/3; 80/3; 117/2; 144/1; 200/1; 265/1; 294/1; 330/3
C;Superfamily: acetylcholine receptor
C;Keywords: brain; glycoprotein; ion channel; neurotransmitter receptor; phosphoprotein;
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-502/Product: nicotinic acetylcholine receptor alpha-7 chain, neuronal #status predi
F;231-254/Domain: transmembrane #status predicted <TR1>
F;262-280/Domain: transmembrane #status predicted <TR2>
F;296-317/Domain: transmembrane #status predicted <TR3>
F;470-488/Domain: transmembrane #status predicted <TR4>
F;46,90,133/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;365,367,413,427,465/Binding site: phosphate (Ser) (covalent) #status predicted
F;415/Binding site: phosphate (Thr) (covalent) #status predicted
F;442/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 47.5%; Score 1253; DB 2; Length 502;
Best Local Similarity 48.8%; Pred. No. 1.1e-98;
Matches 250; Conservative 81; Mismatches 145; Indels 36; Gaps 8;

QY 2 APMALALALLPVSEGGPHEKRLNALLANYNTLPRVANSEPLEVRFGTLTQOIID 61
Db 5 ALMLLAAAGLVRESLQGEFQRLYKELKNYNPLRPVANDSQPLTVYFTLSLMQIMD 64

QY 62 VDEKNQLITNIWLSLEWYNDYLRWDSYGGVKDLRITPNKLMKPDVLMYNSADEGFDG 121
Db 65 VDEKNQVLITNIWLSLEWYNDYLRWDSYGGVKDLRITPNKLMKPDVLMYNSADEGFD 124

QY 122 TYOTNVVRSGGCLYVPPGIFKSTCKMDIAWFPDDOHCMDKFGSWTYDGNOLDVLKXD 181
Db 125 TFEHNVVNSGHCQVLPFGIFKSSCYIDVRWFPDVKCNLKFSGWYGGWLSLQMQE 184

QY 182 EAGGDLSDFTNGEWLIGMPGKNTIYACCPYVDVFTTMRRTLYYFENLIVPC 241
Db 185 --ADISGYISNGEWDLVGPGKTESFYECCKEPYDITFTVTRRTLYYGLNLLIPC 241

QY 242 VLISSSWALLGFTLPPDSGEKLTGLVTLLSLTVFLNLVAETLPQVSDAIPLLGTYNFCIM 301
Db 242 VLISALALLVFLPADSGEKISLIGITVLLSLTVFMLLVAEIMPATSDSVPLIAQYFASTM 301

QY 302 FMVASSVLTIVVLYNHRTADITHMPQWIKSVFLOLWPLILMSRPGKKITRKTMMNT 361
Db 302 IIVGLSVVTVIVLQHHHPDGGKPKWTRVILLNCWAFILMKRPGEDKVRPACQHKQ 361

QY 362 RMRELKERSKS---LLANVLIDDDFR---HGPPPPNSTASTGNLPGCSIFRT 412
Db 362 RRCSSMEMNTVSGQCSNGNMLYI--GFRGLDGVHCTPTTDSGVICGRM--TCS--- 413

QY 413 DFRSFRVPSMTMEDVGGGLGSHH-----RELHLILRELQITARMKKADEAELISDWKF 467
Db 414 -----PTEENL-----LHSGHPSEGDPLAKILEEVRIANFRDQDEEAICNEKF 462

QY 468 AAMVDFCHVFETFTIATVAVLLSAPHII 499
Db 463 AASVDELCLMAFSVFTIICIGILMSAPNFV 494

RESULT 4
G02259
alpha 7 neuronal nicotinic acetylcholine receptor - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 20-Aug-1999
C;Accession: G02259
R;Leonard, S.
submitted to the EMBL Data Library, November 1995
A;Accession: G02259
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-502 <LEO>
A;Cross-references: EMBL:U40583; NID:g1125076; PIDN:AAA83561.1; PID:g1125077
C;Superfamily: acetylcholine receptor

Query Match 47.4%; Score 1252.5; DB 2; Length 502;
Best Local Similarity 48.4%; Pred. No. 1.3e-98;
Matches 249; Conservative 74; Mismatches 140; Indels 51; Gaps 7;

QY 8 LALLA-LLPVSEGGPHEKRLNALLANYNTLPRVANSEPLEVRFGTLTQOIIDVDEKN 66
Db 10 LALAALLHVSLSQGEFQRLYKELKNYNPLRPVANDSQPLTVYFSLNLQIMDVDEKN 69

QY 67 QLLITNIWLSLEWYNDYLRWDSYGGVKDLRITPNKLMKPDVLMYNSADEGDTGYQTN 126
Db 70 QVLTITNIWLSQMTDHYLQWNVSEYGVKTVRFPDGIWKPDILLYNSADERDAFHTN 129

QY 127 VVVRSGGCLYVPPGIFKSTCKMDIAWFPDDOHCMDKFGSWTYDGNOLDVLKDEAGD 186
Db 130 VLVPNSGHCQVLPFGIFKSSCYIDVRWFPDVOHCKLKFSGWSYGGWLSLQMQE---AD 186

QY 187 LSPFITNGEWLIGMPGKNTIYACCPYVDVFTTMRRTLYYFENLIVPCVLIS 246
Db 187 ISGIIFNGEWDLVGIFGKSERFYECCKEPYDITFTVTRRTLYYGLNLLIPCULISA 246

QY 247 MALLGFTLPPDSGEKLTGLVTLLSLTVFLNLVAETLPQVSDAIPLLGTYNFCIMFVAS 306
Db 247 LALLVFLPADSGEKISLIGITVLLSLTVFMLLVAEIMPATSDSVPLIAQYFASTMIIVGL 306

QY 307 SVVLTIVVLYNHRTADITHMPQWIKSVFLOLWPLILMSRPGKKITRKTMMNTMREL 366
Db 307 SVVTVIVLQYHHHPDGGKPKWTRVILLNCWAFILMKRPGEDKVRPACQHKQRCSL 366

QY 367 ELKERSKSLANVLIDDDFRHGPPTNSTAGNL-----GPGC 407
Db 367 ASVEMSAVA-----PPP---ASGNLLYIGRGLDGVHCVPTPDSGVVC 407

QY 408 SIFRTDFRSFRVPSMTMEDV--GGGLGSHHRELHLILRELQITARMKKADEAELISDW 465
Db 408 G-----RMACSPTHDEHLHGGQPEGDPDLAKILEEVRIANFRDQDESEAVCSEW 460

QY 466 KFAAMVDFCLVFTIATVAVLLSAPHII 499
Db 461 KFAACVVDRLCLMAFSVFTIICIGILMSAPNFV 494

RESULT 5
T01378
nicotinic receptor alpha 7 chain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 21-Jul-2000
C;Accession: T01378
R;Seguela, P.; Wadiche, J.; Dineley-Miller, K.; Dani, J.A.; Patrick, J.W.
J. Neurosci. 13, 596-604, 1993
A;Title: Molecular cloning, functional properties, and distribution of rat brain alpha
A;Reference number: Z14310; MUID:93147931; PMID:7678857
A;Accession: T01378
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-502 <SEG>
A;Cross-references: EMBL:S53987; NID:g264770; PIDN:AAB25224.2; PID:g5705903
A;Experimental source: brain
C;Superfamily: acetylcholine receptor

Query Match 47.2%; Score 1246.5; DB 2; Length 502;
Best Local Similarity 49.4%; Pred. No. 4.1e-98;
Matches 248; Conservative 78; Mismatches 149; Indels 27; Gaps 7;

QY 8 LAL-LLPVSEGGPHEKRLNALLANYNTLPRVANSEPLEVRFGTLTQOIIDVDEKN 66
Db 10 LALAALLHVSLSQGEFQRLYKELKNYNPLRPVANDSQPLTVYFSLNLQIMDVDEKN 69

QY 67 QLLITNIWLSLEWYNDYLRWDSYGGVKDLRITPNKLMKPDVLMYNSADEGDTGYQTN 126
Db 70 QVLTITNIWLSQMTDHYLQWNVSEYGVKTVRFPDGIWKPDILLYNSADERDAFHTN 129

QY 127 VVVRSGGCLYVPPGIFKSTCKMDIAWFPDDOHCMDKFGSWTYDGNOLDVLKDEAGD 186

Db 130 VLVNASGHCQYLPFCIFKSSCYDVRWFFPDVQCKLKFGSWSYGGWSLDLQMQE---AD 186
QY 187 LSDFTNGEWYLGMPGKKNITTYACCPPEYVDVFTTMRRTLYFFNLIVPCVLIS 246
Db 187 ISSYIPNGEWDLMGIPGRNKFYECCKPEYPDVFTYTMERRTLYYGLNLLIPCVLISA 246
QY 247 MALLGFTLPDPSGKLTGLVTLLSLVFLNVAETLPQVSDAIPLLGTGYFNCIMFMVAS 306
Db 247 LALLVFLPADSGEKISLGITVLLSLVFLNVAETLPQVSDAIPLLGTGYFNCIMFMVAS 306
QY 307 SVLTVVVLNHYHRTADITHMPQWIKSVFLQWLPWILRMSRPGKKITRKTITMMNTRREL 366
Db 307 SWVTIVILRVHHDPDGGKPKWTRIIILNCAWFLRMKRPGEKVPAQCCHKPRCSL 366
QY 367 ELKERS-----KSLANVLIDDDFR-----HGPPPNSTASTNGLPGCCSIFRTDFRS 417
Db 367 ASVELSAGAPPTNGNLLYI--GFRLEGHMCAPTDPGCVVCGRL--ACSPTHDEHLMH 422
QY 418 FVRPSTMEDVGGGLGSHHRELHLILRELQFITARMKKADEAEALISDWKFAAMVVDPRCL 477
Db 423 GAHPSGDGP-----DLAKILEEVRYIANRRCODESEVICSEWKFAACVVDPLCL 472
QY 478 FVFTLFTIIATVALLSAPHII 499
Db 473 MAFSVFTIICIGILMSAPNFV 494

RESULT 6

S68388
nicotinic acetylcholine receptor alpha-1 chain precursor (clone Ce21) - Caenorhabditis e
C:Species: Caenorhabditis elegans
C>Date: 06-Dec-1996 #sequence_revision 07-Feb-1997 #text_change 20-Aug-1999
C:Accession: S68588; S57496
R:Ballivet, M.; Alliod, C.; Bertrand, S.; Bertrand, D.
J. Mol. Biol. 259, 261-269, 1996
A:Title: Nicotinic acetylcholine receptors in the nematode *Caenorhabditis elegans*.
A:Reference number: S68587; MUID:96196478; PMID:8627624
A:Accession: S68588
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-498 <BAL>
A:Cross-references: EMBL:X83887; NID:9872087; PIDN:CAA58764.1; PID:9872088
C:Superfamily: acetylcholine receptor
C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membrane;
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-498/Product: nicotinic acetylcholine receptor alpha-1 chain #status predicted <MAT>

Query Match 46.9%; Score 1237.5; DB 2; Length 498;
Best Local Similarity 47.5%; Pred. No. 2.4e-97;
Matches 245; Conservative 80; Mismatches 140; Indels 51; Gaps 9;

QY 2 APLAALLALLPVSEQGPHEKRLNALLANNYTLERPVAESEPLEVRFGLTLQIID 61
Db 15 APTLGSL-----QERLYEDLMRYNNLRPVAHSEPVTVHLKVALQIID 61
QY 62 VDEKNQLLITNIWLSLEWNYLNRWDSYEGGVKDLRITPNKLPKDPVLMYNSADEGFDG 121
Db 62 VDEKNQVVVYNAWLDYTWNDYLVWDKAEYGNITDVRFPAGKIWKPDVLLNYSVDTNFD 121
QY 122 TYQTNVVRSGGSLYVPPGIFKSTCKMDIAWFFPDQCHDKMGKSWTYDGNQLDLVLKD 181
Db 122 TYQTNMIVYSGLVVHWPVPGIFKLSCKIDIQWFFPDQCFKFGSWTYDGYKLD--LQP 179
QY 182 EAGG-DLSDFTNGEWYLGMPGKKNITTYACCPPEYVDVFTTMRRTLYFFNLIVP 240
Db 180 ATGFGDISEYISNENWALPLTTVERNEKFDCCPEPDPVHFYLUHMRRTLYYGFNLMP 239
QY 241 CVLISSMALLGFTLPDPSGKLTGLVTLLSLVFLNVAETLPQVSDAIPLLGTGYFNCI 300
Db 240 CILATLMTLLGFTLPDAGEKITQITVLLSICFFLSIVSEMSPTSEAVPLLGIFTCC 299
QY 301 MFMVASSVLTVVVLNHYHRTADITHMPQWIKSVFLQWLPWILRMSRPGKKITRKTII-M 358

Db 300 MIVVTASTVFTVYVNLNHYRTPETHDMPWTRNLLLYWILRMKRGHNLTYASLP 359
QY 359 MNTMRRELEUKERSKSLANVLID-----IDDFRKGPPPNSTASTNGLPGPCS 408
Db 360 PSTK-----PNRHSSEILRNKONEHSLSRANGFDADCRNLQVIMTQSVNGLTSLG-S 412
QY 409 IFRDTFRSFRVPRTMEDVGGGLGSHHRE-----LHLILRELQFITARMKKADEAEALIS 463
Db 413 I-----PSTMISSNGTTTDSQAATLLILHRYHELKIVTRKMIEGDKBEQACN 461
QY 464 DWKEFAAMVVDPRCLFVFTLFTIIATVALLSAPHII 499
Db 462 NWKFAAMVVDRLCLVYFTIIVTIGIFWSPAPLV 497

RESULT 7

JH0173
alpha-bungarotoxin-binding protein alpha-2 chain precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 20-Aug-1999
C:Accession: JH0173
R:Schoepfer, R.; Conroy, W.G.; Whiting, P.; Gore, M.; Lindstrom, J.
Neuron 5, 35-48, 1990
A:Title: Brain alpha-bungarotoxin binding protein cDNAs and WABs reveal subtypes of thi
A:Reference number: JH0172; MUID:90315158; PMID:2369519
A:Accession: JH0173
A:Molecule type: mRNA
A:Residues: 1-511 <SCH>
A:Cross-references: GB:X52296; NID:G63081; PIDN:CAA36544.1; PID:G63082
A:Experimental source: brain
A:Note: this sequence is similar to acetylcholine receptor alpha chains
C:Comment: Alpha-bungarotoxin binding proteins are localized to extrasynaptic pseudoden
C:Superfamily: acetylcholine receptor
C:Keywords: glycoprotein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-510/Product: alpha-bungarotoxin binding protein alpha-2 chain #status predicted <AI
F:239-262/Domain: transmembrane #status predicted <TM1>
F:270-288/Domain: transmembrane #status predicted <TM2>
F:304-323/Domain: transmembrane #status predicted <TM3>
F:479-496/Domain: transmembrane #status predicted <TM4>
F:54/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 46.5%; Score 1226.5; DB 2; Length 511;
Best Local Similarity 46.6%; Pred. No. 2.1e-96;
Matches 246; Conservative 75; Mismatches 134; Indels 73; Gaps 9;

QY 5 LAALALLALLPVSEQGPHEKRLNALLANNYTLERPVAESEPLEVRFGLTLQIIDVDE 64
Db 16 LWASLFLSFFKVSQGGESQRRLYRDLRLNRYLRLRPVWNSQPIVVELQLSLQIIDVDE 75
QY 65 KNQLLITNIWLSLEWNYLNRWDSYEGGVKDLRITPNKLPKDPVLMYNSADEGFDGYQ 124
Db 76 KNQVLIITNAWLQMTWVDIYLSWDQYEGPVQVQNLRFPSDQIWPVDPILLNYSADERFATFH 135
QY 125 TNVVRSGGSLYVPPGIFKSTCKMDIAWFFPDQCHDKMGKSWTYDGNQLDLVLKDEAG 184
Db 136 TNVLVNYSGSQYTPGILKSTCYIDVRWFFPDVQKCDLKEGWSHGLIDQLMLE--- 192
QY 185 GDLSDFITNGEWYLGMPGKKNITTYACCPPEYVDVFTTMRRTLYFFNLIVP 244
Db 193 ADISNYISNGEWDLVGVGPKRNLRYECCKPEYPDVFTTMRRTLYYGLNLLIPCVLI 252
QY 245 SSALLGFTLPDPSGKLTGLVTLLSLVFLNVAETLPQVSDAIPLLGTGYFNCIMFV 304
Db 253 SGIALVFLPADSGEKISLGITVLLSLVFLNVAETLPQVSDAIPLLGTGYFNCIMFV 312
QY 305 ASSVLTVVVLNHYHRTADITHMPQWIKSVFLQWLPWILRMSRPGKKITRKTITMMNTRM 364
Db 313 GLSVVTVLVLFQHHDPDQAGKMPWRVVRVILLNCAWFLRMKKEGNI-----K 361
QY 365 ELEIK-----ERSKSLANVLIDDDFRHGGPPPNSTASTNGL-----GPGCSIF 410

Db 362 PLSCYSYKPKHPSLKNTEENVL-----PQHQPNGNMIYSYHTMENPCC--- 406

QY 411 RTDFRSFVRPSTMEDVGGGLG-----SHRELHL-----ILRELQFITAR 451

Db 407 -----PQN-NDLGSKSGRITCPLSDNEHVKQKALMDTIPVIVKILEBVQFTAMR 455

QY 452 MKKADAEALISDWKFAAMVVDVFLVFTLTIATVALLGAPHII 499

Db 456 FRKQDEGEICSEKFAAAVIDRLCLVAFTLFAICTFTILMSAPNFI 503

RESULT 8

T25671

hypothetical protein D2092.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C;Accession: T25671

R;Gattung, S.; Maggi, L.

submitted to the EMBL Data Library, February 1997

A;Description: The sequence of C. elegans cosmid D2092.

A;Reference number: Z20067

A;Accession: T25671

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-461 <GAT>

A;Cross-references: EMBL:U88167; PIDN: AAB42223.1; GSPDB: GN00019; CESP: D2092.3

A;Experimental source: strain Bristol N2; clone D2092

C;Genetics:

A;Gene: CESP: D2092.3

A;Map position: 1

A;Introns: 36/1; 80/3; 119/2; 191/1; 243/1; 274/1; 303/1; 382/2

C;Superfamily: acetylcholine receptor

Query Match 42.9%; Score 1132; DB 2; Length 461;

Best Local Similarity 44.7%; Pred. No. 2.1e-88;

Matches 216; Conservative 86; Mismatches 127; Indels 54; Gaps 6;

QY 23 EKLLNALLANYTLERPVANSEPLEVRFGTLTQQIIDVDKKNOLLTIWISLEWNY 82

Db 26 ETKLFTDLLKGYPLERPVQNSQPLVKLFLQQLDLDVDEKKNQIVSVNANLSYTFDH 85

QY 83 NLRWNSEYGGVLDLRI--TPNKLKPDVLMYNSADGFGDTQYTNVVRSGSGCLVVP 140

Db 86 KLQWPKYGGIQDIRPFGSSDHIWKPDVLLYNSAEDFSTFKSNLLTHTGTVWIPP 145

QY 141 GIFKSTCKMDIAWPPFDQDCHDKMFGSWTYDGNQLDLVLKDEAGG----DLSDFITNGEW 196

Db 146 GVLKFCQLDVTWPPFDQDVCENKFGSWTFHGYAIDLQIDDDTNGTQSMDLSTVLNGEW 205

QY 197 YLIGMPGKNTIYACCPPEYVDVTIMIRRTLYFFNLIVPCVLISMAILGFTLPP 256

Db 206 QVLTSTNAKRVSYKCCPEPYTVNYLHTRRTLYYGFNLITPSLLISMAILGFMFP 265

QY 257 DSGEKLTGVTLLSLVFLNLVAETLPQVSDAIPILGTYFNCIMFVASSVLTIVVLN 316

Db 266 DAGEKITLVTLLAIVFLISMVSEMTPTTSEAVPLIGVFSCMLVVSASVFTIVLN 325

QY 317 YHRTADIHEMPQIKSVFLOLTPWILMGRPGKTIKTKTMMTMRELEKRSKSL 376

Db 326 LHFRSADSHENPLVRVLEFLPWLFLMGRPGKYFKV----- 363

QY 377 LANVLIDDDFRGPPPNSTASTGNLPGCSIFRTDFRSFVRPSTMEDVGGGLSHHR 436

Db 364 -ANVIDSTKMKPKNPLD-----CNL-----PSN-----HAGYEAOIL 397

QY 437 ELHLIRELOFITARMKKADEAEELISDWKFAAMVVDVFLVFTLTIATVALLSAP 496

Db 398 LLHSVHTELRRVAFYKKEEDRIQTDWFAAMVVDVACLLFTVIVISILAIMMSAP 457

QY 497 HII 499

Db 458 HII 460

RESULT 9

T19862

hypothetical protein C40C9.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C;Accession: T19862

R;Hembry, C.

submitted to the EMBL Data Library, March 1996

A;Reference number: Z19188

A;Accession: T19862

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-542 <WIL>

A;Cross-references: EMBL:Z70286; PIDN: CAA94206.1; GSPDB: GN00028; CESP: C40C9.2

A;Experimental source: clone C40C9

C;Genetics:

A;Gene: CESP: C40C9.2

A;Map position: X

A;Introns: 11/3; 69/3; 123/2; 173/3; 216/1; 248/1; 279/1; 314/1; 350/3; 430/1; 464/1; 4

C;Superfamily: acetylcholine receptor

Query Match 41.0%; Score 1083.5; DB 2; Length 542;

Best Local Similarity 42.6%; Pred. No. 3.5e-84;

Matches 232; Conservative 86; Mismatches 152; Indels 75; Gaps 13;

QY 10 LLALLPVSEOGPHEKILLNALLANYTLERPVANSEPLEVRFGTLTQQIIDVDKKNOLL 69

Db 17 LFHLTEVHSSADEYRLLADLRHNDYPERPVANSEPLEVSVKIVLQQLLDVDEKNQVI 76

QY 70 ITNIMLSLEWNYNLRWNSEYGGVLDLRI--TPNKLKPDVLMYNSADGFGDTQYTNV 127

Db 77 TLVAMIEYQMTDYKLKWDSEYGGIKDIRIPGNANAIWKPDVLLYNSADNFSTTPVNY 136

QY 128 VVRSGSGCLVPPGIFKSTCKMDIAWPPFDQDCHDKMFGSWTYDGNQLDLVLKDEAGDL 187

Db 137 VVSITGVDLQVPPGILLKLSCKIDITYFPDDQDCHLKFSGSWTYSGNFIDLRINGPEGKNI 196

QY 188 SD-----FITNGEWYLGMPGKNTIYACCPPEYVDVTIMIRRTLYFFNLIVP 240

Db 197 SDEGIDVQYVQNGEWNLLAVPARHETNIFD--EQYPYSLFFYLIITQRTLYYGLNLIIP 254

QY 241 CVLISMAILGFTLPPDSEKLTGVTLLSLVFLNLVAETLPQVSDAIPIL-----G 294

Db 255 SFILISMTVLGFTLPPDAGEKITLITILLVSCFFLSWADMTPTTSEAVPLIGLIIFSG 314

QY 295 TYFNCIMFVASSVLTIVVLNYHRTADIHEMPQIKSVFLOLTPWILMGRPGKTIKTR 354

Db 315 AFFSCMLVVSASVFTIVLNLRKPETHEMSPPFLRELLLIWPLLLMRPG----- 369

QY 355 KTIMMTRMRELEKRSKSLANVLIDDDFRHGPFPNPNSTASTGNLGP--GCSIFRT 412

Db 370 KTI FNCTHLKAEAEKAKQGSIKNGV-----GPGKP-----TDSVHPSEGLSLMKN 416

QY 413 -----DFRSP-----VRPSTM-----ED-VGGGIGSHHREL 438

Db 417 IKLGRQOTIDFEYEFHVQHNLMPVAPSEMTPTVTVYSKVNASEYVDVVMTELNKMQKA 476

QY 439 HLILRELQFITARMKKADEAE-----LIISDWKFAAMVVDVFLVFTLTIATVALLS 494

Db 477 CLELKNISSQTRAMRKWEEDERDEQAANDKFAAMVVDVRCCLITFTSVFVIVSTCGIMES 536

QY 495 APHII 499

Db 537 SPHLI 541

RESULT 10

T19622

hypothetical protein C3IH5.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C;Accession: T19622

R;Kershaw, J.
Submitted to the EMBL Data Library, April 1997
A;Reference number: Z19153
A;Accession: T19622
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-560 <WTL>
A;Cross-references: EMBL:Z93778; PIDN:CAB07843.1; GSDB:CN00019; CESP:C31H5.3
A;Experimental source: clone C31H5
C;Genetics:
A;Gene: CESP:C31H5.3
A;Map position: 1
A;Insertions: 24/1; 70/3; 139/2; 270/1; 299/1; 336/3; 372/2; 456/3
C;Superfamily: acetylcholine receptor

```

Query Match      40.9%; Score 1081; DB 2; Length 560;
Best Local Similarity 39.4%; Pred. No. 6e-84;
Matches 218; Conservative 88; Mismatches 155; Indels 92; Gaps 8;

Qy      20 GPHEKRLINALLANVTILERVANESBLEYRFGTLQOIIDVDEKNQLLTINILSLEW 79
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      80 NDYNLRWNSDYGYGKOLIRITFNKLWKDVLIMYNSADEGFGDTQTNNVVRASGGSCILYVP 139
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      88 TDHYLTWDPSFEGNIKEVRLPINNIWKFDVLLYNSVDQOFDSWPVNAVVLTYGTNVITWP 147
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      140 PGIKFSCKMDIANFPDDQDCHDMKFGSWTVDGNQLDLVLKDEAGGLSDFITNGEWLI 199
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      148 PAIRSRCAIDIAYPFDQTQCTMKFGSWTYSGFDTLI---NTTISPATYKPNGEWELL 204
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      200 GMPCKKNTIYACCPPEYVDVTFITMIRRRRLTYFFENLIVPCVLISMAILGFTLPDPSG 259
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      205 GLTSQRSIFFECCPEPYDVTFFVSTRRLTYGFNLLPCMLISLALLSFTLPDCCG 264
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      260 EKLTLGVITLLSLTVFLNVAETLPQVSDAIPLLGTGFNCIMFMVASSVLLTVVVLNYHH 319
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      265 EKNLGVITFMSLVCVFMIAVEAMPQTSDALPLIQIYFSCIMFQVGASVATVIALNFHH 324
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      320 RTADIHE-MPQWIKSVFLOMLPWILMRSP-----GKKITREKTIMMNTMR 364
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      325 RSPEQYKPMNFKLTLLGLWFLTLGMRPDVLSVGHAGHYASDNKKQXQYLIEVERH 384
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      365 ELELKERSKSLLANVLIDID-----DDFRHGPPPNSTAST--GNLGPCCSIFRTDFR 415
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      385 ILTRPNCNGHSVDKAVHLDLSTGNPHSDAKKSSPSPKRTSASIMWTG----- 433
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      416 RSFVRPSTMEDVGGGLGS----- 433
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      434 ----LPTT--QMNGALDSSINKYCTCTKVTRPLENGSATINHKSSPQININNINNYKCAN 487
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      434 ----HREHLHLIRLQFTIARMKADEAEALISDKKAAAMVVDVFCLFVFTLFIIA 487
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      488 NQKTFQEDRHPHHILNELRVLISARVKEEAHMLQADMFASRVVDVRCFLAFSAFLPMC 547
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      488 TVAVLLSAPHIIV 500
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Qy      548 TAIISYNAHLFV 560
      Db      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

```

RESULT 11

A53956
nicotinic acetylcholine receptor alpha-3 chain precursor, neuronal - human
A53956
C:Species: Homo sapiens (man)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 20-Aug-1999
C:Accession: A53956; S21338
R:Mihoilovic, M.; Roses, A.D.
Exp. Neurol. 111, 175-180, 1991
A:Title: Expression of mRNAs in human thymus coding for the alpha3 subunit of a neuronal
A:Reference number: A53956; MUID:9114756; PMID:1989896
A:Accession: A53956
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-503 <MI>
A:Cross-references: GB:W37981; NID:g189252; PIDN:AAA59942.1; PID:g189253
R:Anand, R.; Lindstrom, J.
Submitted to the EMBL Data Library, June 1990
A:Description: Nucleotide sequence of the mature human nicotinic acetylcholine receptor
A:Reference number: S21338
A:Accession: S21338
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 30-503 <RNA>
A:Cross-references: EMBL:X53559; NID:g34985; PIDN:CAA37625.1; PID:g34986
C:Genetics:
A:Gene: GDB:CHRNA3
A:Cross-references: GDB:125219; OMIM:118503
A:Map position: 15q24-15q24
C:Superfamily: acetylcholine receptor
C:Keywords: neurotransmitter receptor

Query Match	37.5%	Score	989	DB 2	Length	503
Best Local Similarity	38.9%	Pred. No.	3.6e-76			
Matches	202	Conservative	91	Mismatches	164	Indels
	62	Gaps	6			
QY	1	MAPMLAALALLPSEQGPHEKRLNALLANVTLERPVANESEPLEVREGLTLCQOII	60			
Db	12	VAPFL-LLLSLPLPARASEAHRFLERUFEDYDNIIRPVANSDPVIHFVFSWSQLV	70			
QY	61	DVDEKNQLLTINWLSLEWMDYNLRWDSYGGVKDLIRTPNKLWKPDPVLMTYNSADEGFD	120			
Db	71	KVDEVNQIMETNLWLKQIWNVDYKWNPSDYGAEFMRVPAQKIWKDPDVLVYNNVAGDFQ	130			
QY	121	GTQTNVVRSGGSLYVPPGIFKSTCKMDIANFPDDQHQDMKFGSWYVDGNQLDLVLK	180			
Db	131	VDDTKALLKYTGVTWIPPAIFKSSCKIDVTYFFPDYQNTKMGFSWSDYAKIDLVI	190			
QY	181	DEAGGDLSDFITNGEAWLIGMPGKKTIIYACCEPYVDVTFITMIRRTLYYFFNLIVP	240			
Db	191	G-SMNLKDYWSEGEWAIKAPGYKHDIKYNCBEEIYPDITSLYRRFLPFTYINLIIP	249			
QY	241	CVLISSMALIGFTLPDPSGEKLTGVTIILSLTVFLNLVAETLPQVSDAIPILGTTFNCI	300			
Db	250	CLLISFLTVLFFVLPDSOGEKVLCTISVLLSLTVFLLVITETIPSTLSIPLIGEYLLFT	309			
QY	301	MFVASSVLTVVVLNVHRTADIHEMPOWIKSVFLOWLPWILMRSPGKKIIRKTIMMN	360			
Db	310	MIFFTLISIVITFVLNVHVRTPTTHMPWSVKWTFVFLNLLPRVFMETRP	357			
QY	361	TRMRELEKRSKSLLANVLIDDDDFRGCPPPPNSTAGNL-----GPCCS-----	408			
Db	358	-----TSNIGNAQKRPGLYGAELSNLNCFSRAESGCKEGYPC	395			
QY	409	-----IFRTDFRRSFVRPSTMEDVGG--GLGSHHRELHLILRELOFITARMKK	454			
Db	396	QDGMCGYCHHRRIKIENFSANLTRSSSSVSDAVLSLSALSPEIKAIOSVKYIAENMKA	455			
QY	455	ADAEALISDWKFAAMVDRFCLFVFTLFIATVALL	493			
Db	456	ONEAKIIDDWKVAMVIDRIPLFWVFTLVCILGTAGFL	494			

RESULT 12

A33001 12
 A37040
 nicotinic acetylcholine receptor alpha-3 chain precursor, neuronal (version 2) - human
 C:Species: Homo sapiens (man)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 21-Jul-2000
 C:Accession: A37040; S24595
 R:Fornasari, D.; Chini, B.; Tarroni, P.; Clementi, F.
 Neurosci. Lett. 111, 351-356, 1990
 A:Title: Molecular cloning of human neuronal nicotinic receptor alpha-3-subunit.
 A:Reference number: A37040; MUID:90245296; PMID:2336208
 A:Accession: A37040
 A:Molecule type: mRNA
 A:Residues: 1-502 <FOR>
 A:Cross-references: EMBL:X52239; NID:q177897; PID:AAC84176.1; PID:q177898

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:18:28 ; Search time 9.35654 Seconds
(without alignments)
2788.120 Million cell updates/sec

Title: US-09-303-232-6
Perfect score: 2640
Sequence: 1 MAPMLAALALLPVSSEQ.....LFTIATVALLSAPHIIQ 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1258.5	47.7	502	1 ACH7_HUMAN	P36544 homo sapien
2	1255.5	47.6	502	1 ACH7_MOUSE	P49582 mus musculu
3	1253	47.5	502	1 ACH7_CHICK	P22770 gallus gall
4	1246.5	47.2	502	1 ACH7_RAT	Q05941 rattus norv
5	1240.5	47.0	499	1 ACH7_BOVIN	P54131 bos taurus
6	1237.5	46.9	498	1 ACH1_CAEBL	P48180 caenorhabdi
7	989.5	37.5	503	1 ACH3_HUMAN	P32297 homo sapien
8	977.5	37.0	499	1 ACH3_RAT	P04757 rattus norv
9	970.5	36.8	495	1 ACH3_BOVIN	Q07263 bos taurus
10	964.5	36.5	567	1 ACH1_DROME	P09478 drosophila
11	948.5	35.9	496	1 ACH3_CHICK	P09481 gallus gall
12	946.5	35.9	496	1 ACH1_MANSE	P91766 manduca sex
13	944	35.8	557	1 ACH1_SCHGR	P23414 drosiocerc
14	943	35.7	576	1 ACH2_DROME	P17644 drosophila
15	933	35.3	494	1 ACH6_HUMAN	Q15825 homo sapien
16	924.5	35.0	528	1 ACH2_CHICK	P09480 gallus gall
17	922.5	34.9	529	1 ACH2_HUMAN	Q15822 homo sapien
18	919.5	34.8	512	1 ACH3_CARAU	P18845 carassius a
19	917	34.7	511	1 ACH2_RAT	P12389 rattus norv
20	915.5	34.7	494	1 ACH6_CHICK	P49581 gallus gall
21	914.5	34.6	493	1 ACH6_RAT	P43143 rattus norv
22	910.5	34.5	495	1 ACHP_RAT	P12392 rattus norv
23	908.5	34.4	519	1 ACHA_DROME	P25162 drosophila
24	906	34.3	457	1 ACHA_BOVIN	P02709 bos taurus
25	900	34.1	521	1 ACH3_DROME	P04755 drosophila
26	899.5	34.1	500	1 ACHN_RAT	P12390 rattus norv
27	897.5	34.0	627	1 ACH4_HUMAN	P43681 homo sapien
28	895	33.9	456	1 ACHA_CHICK	P09479 gallus gall
29	895	33.9	502	1 ACHN_HUMAN	P17787 homo sapien
30	891	33.8	622	1 ACH4_CHICK	P09482 gallus gall
31	890.5	33.7	629	1 ACH4_MOUSE	O70174 mus musculu
32	890	33.7	457	1 ACHA_RAT	P25108 rattus norv
33	890	33.7	470	1 ACHP_CHICK	P26153 gallus gall

RESULT 1

ID	ACH7_HUMAN	STANDARD;	PRT;	502 AA.
AC	P36544; Q15826; Q96RH2; Q99555; Q9BXH0;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Neuronal acetylcholine receptor protein, alpha-7 chain precursor.			
GN	CHRNA7 OR NACHRA7.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=94195283; PubMed=8145738;			
RA	Peng X., Katz M., Gerzanich V., Anand R., Lindstrom J.;			
RT	"Human alpha 7 acetylcholine receptor: cloning of the alpha 7 subunit from the SH-SY5Y cell line and determination of pharmacological properties of native receptors and functional alpha 7 homomers expressed in Xenopus oocytes.";			
RT	Mol. Pharmacol. 45:546-554(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Hippocampus;			
RA	Logel J., Drebing C., Barnhart M., Antle C., Leonard S.;			
RL	Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97062879; PubMed=8906617;			
RA	Elliot K.J., Ellis S.B., Berckhan K.J., Urrutia A.,			
RA	Chavez-Noriega L.E., Johnson E.C., Velicelebi G., Harpold M.M.;			
RT	"Comparative structure of human neuronal alpha 2-alpha 7 and beta 2-beta 4 nicotinic acetylcholine receptor subunits and functional expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and beta 4 subunits.";			
RL	J. Mol. Neurosci. 7:217-228(1996).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97162233; PubMed=9009220;			
RA	Groot Kormelink P.J., Luyten W.H.M.I.;			
RT	"Cloning and sequence of full-length cDNAs encoding the human neuronal nicotinic acetylcholine receptor (nAChR) subunits beta3 and beta4 and expression of seven nAChR subunits in the human neuroblastoma cell line SH-SY5Y and/or IMR-32.";			
RL	FEBS Lett. 400:309-314(1997).			
RN	[5]			
RP	REVISIONS.			
RA	Groot Kormelink P.J., Luyten W.H.M.I.;			
RL	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Epidermal Keratinocytes;			
RA	Arredondo J., Grando S.A.;			
RT	"Cloning cholinergic receptors in human keratinocytes.";			

P30926 homo sapien
P04756 mus musculu
P19370 carassius a
P02708 homo sapien
P02711 torpedo mar
P09483 rattus norv
P09484 gallus gall
P02710 torpedo cal
P05377 xenopus lae
P22456 xenopus lae
P45963 caenorhabdi
Q98880 brachydanio

ALIGNMENTS

Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE OF 17-502 FROM N.A.
TISSUE=Brain;
RA Doucette-Stamm L., Monteggia L.M., Donnelly-Roberts D., Wang M.T.,
Lee J., Tian J., Giordano T.;
RT "Cloning and sequence of the human $\alpha 7$ nicotinic acetylcholine
receptor.";
EL Drug Dev. Res. 30:252-256(1993).
RN (8)
RP SEQUENCE OF 24-502 FROM N.A.
TISSUE=Retina;
RX MEDLINE=94245214; PubMed=8188270;
RA Chini B., Raimondi E., Elgoyhen A.B., Moralli D., Balzaretto M.,
Heinenann S.F.;
RT "Molecular cloning and chromosomal localization of the human $\alpha 7$ -
nicotinic receptor subunit gene (CHRNA7).";
RL Genomics 19:379-381(1994).
RN (9)
RP SEQUENCE OF 118-129 FROM N.A.
MEDLINE=21818878; PubMed=11829490;
RA Riley B., Williamson M., Collier D., Wilkie H., Makoff A.;
RT "A 3-Mb map of a large segmental duplication overlapping the $\alpha 7$ -
nicotinic acetylcholine receptor gene (CHRNA7) at human 15q13-q14.";
RL Genomics 79:197-209(2002).
RN (10)
RP MASS SPECTROMETRY.
TISSUE=Breast cancer;
RX MEDLINE=21829512; PubMed=11840567;
RA Harris R.A., Yang A., Stein R.C., Lucy K., Brusten L., Herath A.,
Parekh R., Waterfield M.D., O'Hare M.J., Neville M.A., Page M.J.,
Zvelebil M.J.;
RT "Cluster analysis of an extensive human breast cancer cell line
protein expression map database.";
RL Proteomics 2:212-223(2002).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
extensive change in conformation that affects all subunits and
leads to opening of an ion-conducting channel across the plasma
membrane.
CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-
bungarotoxin. The structure is probably pentameric (By
similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- MASS SPECTROMETRY: MW=54157.68; METHOD=MALDI.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X70297; CAA49778.1; -;
CC EMBL; U40583; AAA83561.1; -;
CC EMBL; U62436; AAB40114.1; -;
CC EMBL; Y08420; CAA69697.1; -;
CC EMBL; AF385585; AAK68111.1; -;
CC EMBL; L25827; -; NOT ANNOTATED CDS.
CC EMBL; Z23141; CAA80672.1; -;
CC EMBL; AF323758; AAK19515.1; -;
CC PIR; G02259; G02259.
CC PIR; I37185; ACHUA7.
CC Genew; HGNC:1960; CHRNA7.
CC MIM; 118511; -;
CC GO; GO:0005892; C:nicotinic acetylcholine-gated receptor-chan. .; TAS.
CC GO; GO:0015464; F:acetylcholine receptor activity; TAS.
CC GO; GO:0004889; F:nicotinic acetylcholine-activated cation-se. .; TAS.
CC GO; GO:0000187; P:activation of MAPK; TAS.
CC GO; GO:0006810; P:transport; TAS.
CC InterPro; IPR006029; Neur_chan_IIB.
CC InterPro; IPR006202; Neur_chan_IIB.

DR InterPro; IPR006201; Neur_chan_IIB.
DR Pfam; PF02931; Neur_chan_IIB; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMS; TIGR00860; LIC; 1.
DR PROSITE; PS0236; NEUOTR_ION_CHANNEL; 1.
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 502 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
ALPHA-7 CHAIN.
FT DOMAIN 23 230 POTENTIAL.
FT TRANSMEM 231 255 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 262 280 POTENTIAL.
FT TRANSMEM 296 317 POTENTIAL.
FT DOMAIN 318 469 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 470 490 POTENTIAL.
FT DISULFID 150 164 BY SIMILARITY.
FT DISULFID 212 213 ASSOCIATED WITH RECEPTOR ACTIVATION
(BY SIMILARITY).
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).
FT TRANSMEM 133 133 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 11 11 A -> G (IN REF. 1 AND 7).
FT CONFLICT 58 58 S -> N (IN REF. 2 AND 6).
FT CONFLICT 134 134 S -> P (IN REF. 2 AND 6).
FT CONFLICT 364 364 C -> S (IN REF. 8).
FT CONFLICT 375 375 A -> G (IN REF. 1).
FT CONFLICT 409 413 RMACS -> AMPAP (IN REF. 8).
SQ SEQUENCE 502 AA; 56449 MW; D94B3A482EAA0E42 CRC64;
Query Match 47.7%; Score 1258.5; DB 1; Length 502;
Best Local Similarity 48.6%; Pred. No. 6.3e-102; Indels 51; Gaps 7;
Matches 250; Conservative 75; Mismatches 138;
Qy 8 LALLA-LLPVSEQGFHEKRLNALLANYTLERFVANESEPLEVRFGLTLQOIIDVDKDN 66
Db 10 LALAALLHVSLSQGFQKLYELVKNYNPLRPVANDSQPLTVVFSLSLIQIMVDKDN 69
Qy 67 QLLTNTIWSLEWYNDYLRNDSEYGVKDLRITPNKLVKDPVLMYNSADSGFGTYGTN 126
Db 70 QVLTNTIWLQMSWTDHYLQNVSEYGVKTVRFDPGQIKPDIILYNSADRFATFTN 129
Qy 127 VVVRSGSCLVVPPGIFKSTCKMDIAWFPDDQCHDMKFGSWTYDGNQLDLVLKDEAGD 186
Db 130 VLVNSSHCQVLPGLPKSSCYIDVRWFPDPVQCKLKFGSWSGWSLDLQMBE--AD 186
Qy 187 LSDFITNGEWYLGMPGKKNITTYACCPYVDVFTTITMIRRTLYYFFNLIVPCVLIS 246
Db 187 ISGVIPGEWDLVGPGRSERFEYCKPEYDPDVTFTVMRRRTLYYGLNLLIPCVLISA 246
Qy 247 MALLGFTLPDPSGKLTGLVITLLSLTVFLNVAETLPVSDATPLIGTYENCIMFVVAS 306
Db 247 LALLVFLPADSGEKISLIGITVLSLTVFLNVAEIMPATSDSVFLTAQYFASFTMIIVGL 306
Qy 307 SVLTVVVLNVYHRTADTHEMPQIKSVFLQWLFWILRMSRPGKKITRKTMMTRREL 366
Db 307 SVVTVIVLQYHHDDPGGKMPKWTRVILLNWCWFLRMKPGEDKVRPACQHKQRCSL 366
Qy 367 ELKERSKSLANLVLDIDDDFRHGPPPNPNSTASTGNL-----GPGC 407
Db 367 ASVEMSAVA-----PPP---ASGNLLYIGRGLDGVHCVPTPDGSGVVC 407
Qy 408 SIFTFDRSFVRPSTMEDV--GGGLGSHHRELHLILRELQFTIARMKKADEAEALLSDW 465
Db 408 G-----RMACSTHDEHLHGQPEGDPDLAKILEEVRYIANRFRQDESEAVCSEW 460
Qy 466 KFAAMVDRFCLFVFTLTFTIATVAVLLSAPHII 499
Db 461 KFAACVDRCLMAFSVFTICTIGILMSAFNV 494

RESULT 2

```

ACH7_MOUSE
ID ACH7_MOUSE STANDARD; PRT; 502 AA.
AC P49582;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.
GN CHRNA7 OR ACR47.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95324936; PubMed=7601470;
RX "Urtreger A., Seldin M.F., Baldini A., Beaudet A.L.;
RT "Cloning and mapping of the mouse alpha 7-neuronal nicotinic
RT acetylcholine receptor.";
RL Genomics 26:399-402(1995).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-
CC bungarotoxin. The structure is probably pentameric (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L37663; AAC2053.1; -.
DR PIR; A57175; A57175.
DR MGD; MG1:99779; Chrna7.
DR InterPro; IPR006029; Neu_channel_memb.
DR InterPro; IPR006202; Neu_chan_LBD.
DR InterPro; IPR006201; Neu_channel.
DR Pfam; PF02931; Neu_chan_LBD; 1.
DR Pfam; PF02932; Neu_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRfams; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 502 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
FT ALPHA-7 CHAIN.
FT DOMAIN 23 230 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 231 255 POTENTIAL.
FT TRANSMEM 262 280 POTENTIAL.
FT TRANSMEM 296 317 POTENTIAL.
FT DOMAIN 318 469 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 470 490 POTENTIAL.
FT DISULFID 150 164 BY SIMILARITY.
FT DISULFID 212 213 ASSOCIATED WITH RECEPTOR ACTIVATION
FT (BY SIMILARITY).
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 502 AA; 56631 MW; C9312E5226D120E3 CRC64;

Query Match 47.6%; Score 1255.5; DB 1; Length 502;
Best Local Similarity 49.6%; Pred. No. 1.2e-101;
Matches 249; Conservative 78; Mismatches 148; Indels 27; Gaps 7;

```

QY 8 LAL-LALLPVSQGPHEKRLNALLANYNTERPVANSEPLEVRFLTLTQQIIDVDEN 66

```

Db 10 LALAAALLHVSQGEFORLYKELVKNYPLERPVANDSOPLTVYFSLSLQIMDVDEKN 69
QY 67 QLLITNIWLSLENDYNLRWNSDEYGGVKDLRTTPNKLMPDVMYNSADEGDTGYQTN 126
Db 70 QVLTINILQMSWTDLHYLQWNSEYPGVKNVRPPDQGIWKPDILLYNSADERDAFHTN 129
QY 127 VVVRSGSGLYVPPGIFKSTCKMDIAWFFPDHCHDKMGSWTYDGNQLDLVKDEAGD 186
Db 130 VLNVASGHQYLPPIGFKSCYIDVRWFFDVQOQCKLFGSWSYGGMSLDLQMQE---AD 186
QY 187 LSDFTNGEWYLGMPGKNTITYACCPRYVDVTFTIMIRRTLYFFENLIYPCVLISS 246
Db 187 ISSYIENGWDLGIFGKNEKFECKEYPDVTIVIMRRRLTYGLNLLPCVLISA 246
QY 247 MALLGFTLPDSCGKLTGLVTILLSLTFLNLVAETLPQVSDAIPLLGYFNCIMFMVAS 306
Db 247 LALLVLLPADSGEKISLGTIVLLSLTVFVLLVAEIMPATSDSVPLIAQVFASTMIIVGL 306
QY 307 SVVLTVVVLYNHRTADIHEMPQWIKSVFLOMPLWILRMRSPGKTKTRKIMNTMRREL 366
Db 307 SVVTVIVLYRHHDPDGGKMPKWTIRILLNWCWFLRMKRPGEKVRPACQHKPRCSL 366
QY 367 ELKERSS---KSLLANVLIDDDRR-----HGPPPNSTASTGNLGPSCSIFRTDFRS 417
Db 367 ASVELSAGAGPPTSGNLLYI--GFRLEGHMCAPTPDSGVVCGRL--ACSPHDEHLMH 422
QY 418 FVRPSTMEDVGGGLGSHRHLHLRELQFITARMKKADEAEELISDWKFAAMVVDRCFL 477
Db 423 GTHPSDGP-----DLAKILEEVRYANRRCQDESEVICSEWKEFAACVVDRLCL 472
QY 478 FVPTLTITATVAVLLSAPHII 499
Db 473 MAPSVFTIITIGILMSAPNFV 494
RESULT 3
ACH7_CHICK
ID ACH7_CHICK STANDARD; PRT; 502 AA.
AC P22770;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91097796; PubMed=1702646;
RA Couturier S., Bertrand D., Matter J.M., Hernandez M.C., Bertrand S.,
RA Millar N., Valera S., Barkas T., Ballivet M.;
RT "A neuronal nicotinic acetylcholine receptor subunit (alpha 7) is
RT developmentally regulated and forms a homo-oligomeric channel blocked
RT by alpha-BTX.";
RL Neuron 5:847-856(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90315158; PubMed=2369519;
RA Schoepfer R., Conroy W.G., Whiting P., Gore M., Lindstroem J.;
RT "Brain alpha-bungarotoxin binding protein cDNAs and MAbS reveal
RT subtypes of this branch of the ligand-gated ion channel gene
RT superfamily.";
RL Neuron 5:335-48(1990).
RN [3]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN=White leghorn; TISSUE=Erythrocyte;
RX MEDLINE=93049204; PubMed=1425587;
RA Matter-Sadzinski L., Hernandez M.C., Roztocil T., Ballivet M.,

```

RA Matter J.M.;
 RT "Neuronal specificity of the alpha 7 nicotinic acetylcholine receptor
 RT promoter develops during morphogenesis of the central nervous
 RT system.";
 RL EMBO J. 11:4529-4538(1992).
 [4]
 RN SEQUENCE OF 24-47.
 RP TISSUE=Brain;
 RC MEDLINE=92049732; PubMed=1719423;
 RX MEDLINE=95270494; PubMed=3860855;
 RA Conti-Tronconi B.M., Dunn S.M.J., Barnard E.A., Dolly J.O., Lai F.A.,
 RA Ray N., Raftery M.A.;
 RT "Brain and muscle nicotinic acetylcholine receptors are different but
 RT homologous proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:5208-5212(1985).
 [5]
 RN MUTAGENESIS OF LEU-270.
 RP MEDLINE=93024917; PubMed=1383829;
 RX Recah F., Bertrand D., Galzi J.-L., Devillers-Thiery A., Mulle C.,
 RA Hussy N., Bertrand S., Ballivet M., Changeux J.-P.;
 RT "Mutations in the channel domain alter desensitization of a neuronal
 RT nicotinic receptor";
 RL Nature 353:846-849(1991).
 [6]
 RN MUTAGENESIS TO CONVERT ION SELECTIVITY FROM CATIONIC TO ANIONIC.
 RP MEDLINE=93024917; PubMed=1383829;
 RX Galzi J.-L., Devillers-Thiery A., Hussy N., Bertrand S.,
 RA Changeux J.-P., Bertrand D.;
 RT "Mutations in the channel domain of a neuronal nicotinic receptor
 RT convert ion selectivity from cationic to anionic.";
 RL Nature 359:500-505(1992).
 CC -!- FUNCTION: After binding acetylcholine, the ACHR responds by an
 CC extensive change in conformation that affects all subunits and
 CC leads to opening of an ion-conducting channel across the plasma
 CC membrane.
 CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-
 CC bungarotoxin. The structure is probably pentameric (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- DEVELOPMENTAL STAGE: Alpha-7 transcripts transiently accumulate
 CC in the developing optic tectum between E5 and E16.
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X52295; CAA36543.1; -;
 CC EMBL; X68246; CAA4817.1; -;
 CC EMBL; X68586; CAA48576.1; -;
 CC PIR; JN0113; JN0113.
 CC PDB; 1KC4; 17-APR-02.
 CC PDB; 1KL8; 17-APR-02.
 CC InterPro; IPR006029; Neu channel memb.
 CC InterPro; IPR006202; Neu channel LBD.
 CC InterPro; IPR006201; Neu channel.
 CC Pfam; PF02931; Neu chan LBD; 1.
 CC Pfam; PF02932; Neu chan memb; 1.
 CC PRINTS; PR00252; NRIONCHANNEL.
 CC TIGRfams; TIGR00860; LIC; 1.
 CC PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
 KW Postsynaptic membrane, Ionic channel, Glycoprotein; Signal;
 KW Transmembrane, Multigene family; 3D-structure.
 FT SIGNAL 1 23
 FT CHAIN 24 502 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
 FT ALPHA-7 CHAIN.
 FT DOMAIN 24 230
 FT TRANSMEM 231 255
 FT TRANSMEM 262 280
 FT TRANSMEM 296 317

FT DOMAIN 318 469 CYTOPLASMIC.
 FT TRANSMEM 470 490
 FT DISULFID 150 164
 FT DISULFID 212 213
 FT CARBOHYD 46 46
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MUTAGEN 270 270 L-S,S,T: SUPPRESSES INHIBITION BY THE
 FT OPEN-CHANNEL BLOCKER QX-222.
 FT CONFLICT 26 27 QX -> ET (IN REF. 3).
 SQ SEQUENCE 502 AA; 56946 MW; 572325D4309AD2FD CRC64;
 Query Match 47.5%; Score 1253; DB 1; Length 502;
 Best Local Similarity 48.8%; Pred. No. 1.9e-101;
 Matches 250; Conservative 81; Mismatches 145; Indels 36; Gaps 8;
 Qy 2 APMALALLALLPVSQGEHEKELLNALLANNYNTLPRVANESPELEVRFLTQLOIID 61
 Db 5 ALMLWLLAAAGLVRESLQGEFQKLYKELLKNYPLERPVANDSQPLTVVFTLSLMQIMD 64
 Qy 62 VDEKNQLLTINWLSLEWNYNLRWDSYGGVVDLRTITPNKLVKPDVLMYNSADEGFDG 121
 Db 65 VDEKNQVLTINWLMQYTDHYLQWNVSEYPGVKNVFPDGLIWKPDILLYNSADERFDA 124
 Qy 122 TYQTNVVRSGSCLVYPGGEFKSTCKMDIAWPFDDOHCMDKFGSWTYDGNQLDLVKD 181
 Db 125 TETNVLVNSGHCQYLPGLIFKSSCYDVRWFFDVQKCNKFGSWTYGWSLDLQMOE 184
 Qy 182 EAGGDLSDFTITNGEWYILGMPGKNTITYACCPPEYVDVTFIMIRRTLYFFNLIVPC 241
 Db 185 --ADISSYISNGEWDLVGIPGRKTESFYECCKEYPDITFTVTMRRTLYVGLNLLIPC 241
 Qy 242 VLTISSMALLGFTLPDSDGKLTIGVTILLSITVFLNLVAETLPOVSDAIPLLGTYFCIM 301
 Db 242 VLISALALLVFLPADSGEKISGITVLLSLTVFMLLVAEIMPATSDSVPLIAQYFASTM 301
 Qy 302 FMVASSVVLTVVLYNHHRTADTHEMPQWIKSVFLQWLPILTRMSRPGKKTTRTIMNT 361
 Db 302 IIVGLSVVTVVLYQYHHDPDGGKMPKTRVILLNWCAMFLMRKPGEDKVRPACQHKQ 361
 Qy 362 RMRELEKERSKSKS----LLANVLIDDDDFR-----HGPPPNSTASTGNLGPCCSIFRT 412
 Db 362 RRCSSLMENMTVSGQCCSNGNMLYI--GFGLDGVHCTPTTDSGVICGRM--TCS---- 413
 Qy 413 DFRSRVPRPTMEDVGGGLSHH-----RELHLILRELOFTIARKKADAEALISDWKF 467
 Db 414 -----PTEENL---LHSGHPSEGDPLAKILEVRYIANFRDQDEEAICNEMKF 462
 Qy 468 AAMVVDRLCLFVFTLTFTIATVALLSAPHII 499
 Db 463 AASVVDRLCLMAFSVFIICTIGILMSAPNFV 494
 RESULT 4
 ACH7 RAT
 ID ACH7 RAT STANDARD; PRT; 502 AA.
 AC Q05941;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 21-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.
 GN CHRNA7 OR ACRA7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 EX MEDLINE=93147931; PubMed=7678857;
 RA Seguela P., Wadiche J., Dineley-Miller K., Dani J.A., Patrick J.W.;
 RT "Molecular cloning, functional properties, and distribution of rat

DR TIGRFAMS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 495 NEURONAL ACETYLCHOLINE RECEPTOR PROTHIN,
FT DOMAIN 22 230 ALPHA-3 CHAIN.
FT TRANSMEM 231 255 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 263 281 POTENTIAL.
FT TRANSMEM 296 318 POTENTIAL.
FT TRANSMEM 319 467 CYTOPLASMIC (BY SIMILARITY).
FT DOMAIN 319 487 POTENTIAL.
FT TRANSMEM 468 487 POTENTIAL.
FT DISULFID 149 163 BY SIMILARITY.
FT DISULFID 213 214 ASSOCIATED WITH RECEPTOR ACTIVATION
FT CARBOHYD 45 45 (BY SIMILARITY).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 495 AA; 56914 MW; 322825629821EA07 CRC64;
Query Match 36.8%; Score 970.5; DB 1; Length 495;
Best Local Similarity 39.8%; Pred. No. 8.3e-77;
Matches 198; Conservative 93; Mismatches 181; Indels 25; Gaps 7;
QY 5 LAALALLLPVSEQGPHEKRLNALLANNVTLRPNVASEPLEVFRGLTLOQIIDVDE 64
DB 7 LRLRLLLLPVASTSDAEHRLPERLPEDYNIIRPVANVSDPVIQFEVMSQLVKVDE 66
QY 65 KNQLLITNWLSEWNYLWNSDEYGVKDLRITNKLWKPDLVWVNSADEGFDGTQY 124
DB 67 VQIMETNLWLKQWINDYKLNWNSDYDGAEFMRVPAEKWKPDIYNNNAVGVDFQVDDK 126
QY 125 TNNVVRGGSGLYPPGFKSTKMDIAWPFDDQHCMDKFGSWTYDGNOLDVLKDBAG 184
DB 127 TKALLKTYGVTWIPPAIFKSSCIDYTFYDFYQNTCKMFGSWYDKAKIDVLIG-SS 185
QY 185 GDLSDFTNGEWLIGMGKKNITYACCPPEYVDVFTTMRRTLYFPNLIYPCVLI 244
DB 186 MNLKDYWESGEWAIKAPGYKHDIKNCCEIYDITVSLYIRLPLFYTNLIIPCLLI 245
QY 245 SSMAILGFTLPPSGEKLITGVITLLSLTFLVNLVAETLPOVSDAIIPLGTNYFCIMFW 304
DB 246 SFLTVLVFLYLPSCGGEKVTICISVLLSLTVFLVITETIPSTSLVILIGEYLLFTMIFV 305
QY 305 ASSVVLTVVNLVYHRTADITHMPQWIKSVFLQWLPLMRSPGK---KITRKTIMWNT 361
DB 306 TLSIVITVFLVNLVHYRTPHTMPAWKTFIPLNLLPRVMTFRPASNEGTVQRPFPYSA 365
QY 362 RMRELEKERSKSLLANVLIDDDFRHGPPPPNSTAGNLGPGCSIFR---TDFRRSF 418
DB 366 ELSNLNCFSRIESKVC-----KEGVP-----CQDGLCG-YCHHRRAKISNFSANL 409
QY 419 VRPSTMEDVGG--GLGSHRELHLILRELQFITARMKKADEAEELISDMKFAAMVDRFC 476
DB 410 TRSSSESVDAVLSLSALSPEIKAIQSVKYIAENMKQAQNEAKIQQDKYVAMVIDRIF 469
QY 477 LFVFTLFTIIATVAVLL 493
DB 470 LWFVILVILGTLGFL 486
RESULT 10
ACHI DROME
ID -ACHI DROME STANDARD; PRT; 567 AA.
AC P09478; OSVC74;
DT 01-MAR-1989 (Rel. 10, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Acetylcholine receptor protein, alpha-like chain 1 precursor.
GN NACR-ALPHA-96AA OR ACRB OR ALS OR ACR96AA OR CG5610.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=88283626; PubMed=2840281;
RA Bossy B., Ballivet M., Spierer P.;
RT "Conservation of neural nicotinic acetylcholine receptors from
RL Drosophila to vertebrate central nervous systems.";
EMBO J. 7:611-618(1988).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kechum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclab J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[3]
RN REVISIONS.
RP MEDLINE=22426069; PubMed=12537572;
RX Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC -!- FUNCTION: After binding acetylcholine, the Achr responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: CNS in embryos.
CC -!- DEVELOPMENTAL STAGE: Late embryonic, late pupal and second instar
CC larvae stages.

DR pfam; PF02931; Neur chan LBD; 1.
 DR pfam; PF02932; Neur chan memb; 1.
 DR PRINTS; PR00252; NR1ONCHANEL.
 DR TIGRFAMS; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEURCTR_ION_CHANNEL; 1.
 KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family.
 FT SIGNAL 1 22
 FT CHAIN 23 496
 FT
 FT DOMAIN 23 231
 FT TRANSMEM 232 256
 FT TRANSMEM 264 282
 FT TRANSMEM 298 319
 FT DOMAIN 320 468
 FT TRANSMEM 469 488
 FT DISULFID 150 164
 FT DISULFID 214 215
 FT
 FT CARBOHYD 46 46
 FT CARBOHYD 163 163
 FT SEQUENCE 496 AA; 57027 MW; PD25BC02A9B601FF CRC64;
 Query Match 35.9%; Score 948.5; DB 1; Length 496;
 Best Local Similarity 37.7%; Pred. No. 6.9e-75;
 Matches 195; Conservative 89; Mismatches 168; Indels 65; Gaps 6;
 QY 7 ALALLALLPVSSEQ-----PHEKRLNALLANNTLERPVANSEPLEVRFGLTQQIIDV 62
 Db 6 ALLTLTAACILFQCGGSEPHRIYAALFKNYOFVRPVKNASDFVIFVMSQLVKV 65
 QY 63 DEKNQLATINLWLSLEWYNDYLRWNSYGVGKDLRITPNKLPDVLVLYNSADEGFGDT 122
 Db 66 DEVNQIMETNLWLXHWNDYKLRWNPVDYGGAEIRVPSGQIWKPDIVLYNNAVGDFQVD 125
 QY 123 YQTNVVRSGSGCLYVPPGIPKSTCKMDIAWPPDDQCHDMKFGSWTYDGNQLDLVLKDE 182
 Db 126 DKTKALLKYTGDTWTPPAIFKSSCKIDVTYFPEDYQNCITMKFGSWSDKAKIDLVLIGS 185
 QY 183 AGGDLSPITNGEVLIGMCPGKNITITVACCPPEPVVDVTFIMIRRLTYFFENLIVPCV 242
 Db 186 T-MNLKDYWSEGEWAILKAPGYKHDIKYNCEEYITDITYSLYIRRLPLFTIINMIPCL 244
 QY 243 LISSMALIGFTLPDPSGEKLTGLVTLISLVFLNLVAETLPQVSDAIPLLGTGYFCIMF 302
 Db 245 LISFLTVLVFLPSDQGEKVTLCISVLSLVFLVITETIPSTSLVPLIGEVLLFTMI 304
 QY 303 MVASGVLTVVVLNVHHTADIHEMPQWIKSVFLQWLIPWILRMRPGKKITRKTMMNTR 362
 Db 305 FVTLISIVTVFLNVHVHTPKTHTMPVWVRTIFLNLPRIMFMRP----- 350
 QY 363 MRELELKERSKSLANLVLDIDDDFRHGPPPNSTASGNLG----- 404
 Db 351 -----TSDENNQKPFYTFSEFNSLNCNFSSEIKCKGQFVQCD 390
 QY 405 PGCSIFR-----TDFRSFVRPSTMEDVGG--GLGSHHRELHLILRBLQFITARMKAD 456
 Db 391 MACSCCQYQRMKFSDFSGNLTRSSSESVDPLFSFVLSPEMRDAIESVKVIAENMKQN 450
 QY 457 BEALISDWKFAAMVVRDFCLFVTLFTIATVAVLL 493
 Db 451 EAKGIQDDWKYVAMVIDRIFLWVFLVLCILTAGLFL 487

RESULT 12

ACHI MANSE
 ID ACHI MANSE STANDARD; PRT; 516 AA.
 AC P91766;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acetylcholine receptor protein, alpha-like chain precursor (MARAL).
 GN ARA1.

OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;
 OC Sphingidae; Sphinginae; Manduca.
 OX NCBI_TaxID=71130;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98424077; PubMed=9753155;
 RA Eastham H.M., Lind R.J., Eastlake J.L., Clarke B.S., Townner P.,
 Reynolds S.E., Wolstenholme A.J., Wonnacott S., receptor from the
 FT "Characterization of a nicotinic acetylcholine receptor from the
 RL Insect Manduca sexta.";
 CC Eur. J. Neurosci. 10:879-889 (1998).
 CC -!- FUNCTION: After binding acetylcholine, the ACHR responds by an
 CC extensive change in conformation that affects all subunits and
 CC leads to opening of an ion-conducting channel across the plasma
 CC membrane (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Y0795; CAA70928.1; --
 CC InterPro; IPR006029; Neur channel memb.
 CC InterPro; IPR006022; Neur chan_LBD.
 CC InterPro; IPR006201; Neur channel.
 CC Pfam; PF02931; Neur chan_LBD; 1.
 CC Pfam; PF02932; Neur chan memb; 1.
 CC PRINTS; PR00252; NR1ONCHANEL.
 CC TIGRFAMS; TIGR00860; LIC; 1.
 CC PROSITE; PS00236; NEURCTR_ION_CHANNEL; 1.
 DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family.
 FT SIGNAL 1 21
 FT CHAIN 22 516
 FT
 FT DOMAIN 22 243
 FT TRANSMEM 244 264
 FT TRANSMEM 274 294
 FT TRANSMEM 306 326
 FT DOMAIN 327 465
 FT TRANSMEM 466 486
 FT DISULFID 149 163
 FT DISULFID 222 223
 FT
 FT CARBOHYD 45 45
 FT CARBOHYD 132 132
 FT CARBOHYD 233 233
 FT SEQUENCE 516 AA; 58720 MW; E7A71E8C45D13BD2 CRC64;
 Query Match 35.9%; Score 946.5; DB 1; Length 516;
 Best Local Similarity 39.7%; Pred. No. 1.1e-74;
 Matches 195; Conservative 84; Mismatches 171; Indels 41; Gaps 6;
 QY 21 PHEKRLNALLANNTLERPVANSEPLEVRFGTLQOIIDVDEKNQLLITNWLISLBN 80
 Db 23 PDARKLYDDLLSNYNKLVRFVNLVNSDALTVRAIKLSQLIDVNLKQIMTNTNLWVEQSWY 82
 QY 81 DYNLRWNSDYGVGKDLRITENKLPDVLVLYNSADEGFGDTYQTNVVRSGSGCLYVPP 140
 Db 83 DYKLSWEPREYGGVEMLVHPSDHLWRPDIYLYNNADGNFVTLATKATLNTYGRVETPP 142
 QY 141 GIFKSTCKMDIAWPPDDQCHDMKFGSWTYDGNQLDLVLKDEAG-----DLSDFIT 192
 Db 143 AIYKSSCEIDVEYFPDQTCVMKFGSWTYDGFQVLRHIDVRGTNNVVELGVDLSEFT 202
 QY 193 NGEWYLGMPGKKNTITYACPEPVVDVTFIMIRRLTYFFENLIVFCVLISNALLGF 252

```

Db 203 SVWDILVPAVRNEKYTCDEPYLDITENITMRKILFYVNLIIIPCMGISFLVIVF 262
QY 253 TLPDSEGEKUTLGVITLLSTVFLNVAETLPQVSDAIPLLGYFNCIMFMVASSVLTV 312
Db 263 YLPDSGEKVSLSISLLSTVFLNVAEIIPTSLVPLLGKFLVFTMLDTRFSICVTV 322
QY 313 VVLNYHRTADIHEMPQWIKSVFLOLWPILWMSRPGKKI---TRKTMNTRMRELEIX 369
Db 323 VVLNVHFRSQTHMSPWRRVFIHVLPRLVWRRPHYRDLPHRSRPAGLVTGAGETTLW 382
QY 370 ERSSKILANVLIDDIFRGHPPPPNSTASTGNLGP---GCSIFRTDPRRSFVRPSTMEDV 427
Db 383 DEGSPGV-----PAPRPPPCAPPLAPCAACA-----PAEPAL 416
QY 428 GGGIGSHRR---ELHLILRELQFTARMKKADEABELISDWKFAAMVVDRCLVFILFTI 485
Db 417 CDALRRWHRCPELHKAIDGINIADQTRKEESTRVKEDWKYVAMVLDRLPFLIFTLAVV 476
QY 486 IATVAVILSAP 496
Db 477 VGSAGIILQAP 487

RESULT 13
ACHL_SCHGR
ID ACHL_SCHGR STANDARD; PRT; 557 AA.
AC P23414;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetylcholine receptor protein, alpha-l1 chain precursor.
OS Schistocerca gregaria (Desert locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthoptera; Acridoidea; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
OX NCBI_TaxID=7010;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91092263; PubMed=1702381;
RA Marshall J., Buckingham S.D., Shingai R., Lunt G.G., Goosey M.W.,
RA Darlison M.G., Sattelle D.B., Barnard E.A.;
RT "Sequence and functional expression of a single alpha subunit of an
RT insect nicotinic acetylcholine receptor.";
RL EMBO J. 9:4391-4398(1990).
CC -I- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X55439; CAA39081.1; -.
CC PIR; S12359; S12359.
CC InterPro; IPR006029; Neu_channel_memb.
CC InterPro; IPR006202; Neu_chan_IIBD.
CC InterPro; IPR006201; Neu_channel.
CC Pfam; PF02931; Neu_chan_IIBD; 1.
CC Pfam; PF02932; Neu_chan_memb; 1.
CC PRINTS; PR00252; NEURCHANNEL.
CC TIGRfam; TIGR00860; LIC; 1.
CC PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1 23 POTENTIAL.

```

```

FT CHAIN 24 557 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA-L1
FT CHAIN.
FT DOMAIN 24 244 EXTRACELLULAR.
FT TRANSMEM 245 266
FT TRANSMEM 274 294
FT TRANSMEM 308 329
FT DOMAIN 330 500 CYTOPLASMIC.
FT TRANSMEM 501 523
FT DISULFID 151 165
FT DISULFID 224 225
FT CARBOHYD 47 47
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 382 400 ALA/SER-RICH.
FT DOMAIN 406 422 HIS-RICH.
SQ SEQUENCE 557 AA; 63026 MW; 168389C87DFDF3E CRC64;

Query Match 35.8%; Score 944; DB 1; Length 557;
Best Local Similarity 36.9%; Pred. No. 2e-74;
Matches 201; Conservative 96; Mismatches 181; Indels 66; Gaps 7;

QY 1 MAPMLAALALLPVSEOGPEKRLNLNALLANYNTLREVANESPEVRFGLTQQII 60
Db 5 LPPMLLLLLLLLHPAAANPDAKRLVDLLSNYNRLRPVSNNDTVLVKGLRLSQLI 64
QY 61 DVDEKQQLITNIWLSLEWYNDYLNWNSHYGGVKDLRITPNKLVKDPDVLVMSADGPD 120
Db 65 DLNLKQQLITNVWLEHWDPAHYGGVTELYPSEHILWLPDILVYNNAGEYV 124
QY 121 GTVTNWWVRSGGCLYVPGIFKSTCKMDIAFPDQDQCDKMGKFSWTYDGNQDLIV-- 178
Db 125 VTTMTXAVLHHTGKVWTPPAIFKSCCEIDRVFPDQQTCFMKFGSWTYDGGQIDLKHI 184
QY 179 -----LKDEAGGLSDFITNGEWYLGMPGKNTITYACCPPEYVDVTFIMIRRTLY 232
Db 185 NQKYDDNKVKVIGIDREYVPSVEMDILGVPAERHEKYVPCCAAPYDIFENITLRKTLF 244
QY 233 YFFNLIVPCVLISSMALLGFTLPDSEGEKUTLGVITLLSTVFLNVAETLPQVSDAIP 292
Db 245 YTVNLIVPCVCGISYLSVLVFLPADSGEKIALCISILSQTMFFLLISEIPSTSLAPL 304
QY 293 LGTYFCIMFMVASSVLTVVVNLVYHRTADIHEMPQWIKSVFLOLWPILWMSRPGK-- 350
Db 305 LGXYLLFTVVLGSLVVTIMVNLVHYRKPSTHKMAPVWKVFIKRLPKLLLMRVPEQLL 364
QY 351 -KITRKTIMMNTMRRELKRSKSLLA--NVLDIDDDFRH----- 389
Db 365 ADLASKRLLRHAHNSKLSAAAAAASSSAASPDLSLRHHHLHQHQLHLQLHLQ 424
QY 390 GPPPPNST-ASTGNLPGGCSIFRTDPRRSFVRPSTMEDVGGGLGS----- 433
Db 425 RPGGCGNLHSATNRFSGSAGAF-----GGLPVVGLDGSLDVATR 466
QY 434 -HHRLEHLTLRELQFTARMKKADEABELISDWKFAAMVVDRCLVFILFTIIATVAVL 492
Db 467 KYPFELEKAIHNVLFIONHQRODEDAEDQDWGFVAMVLDRLFLWIFTIASIVGTFAIL 526
QY 493 LSAP 496
Db 527 CEAP 530

RESULT 14
ACH2_DROME
ID ACH2_DROME STANDARD; PRT; 576 AA.
AC P17644; Q9VC73;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Acetylcholine receptor protein, alpha-like chain 2 precursor.
GN NACR-ALPHA-96AB OR ACR96 OR SAD OR ACR96AB OR CG6844.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RX MEDLINE=20301489; PubMed=2114015;
RA Baumann A., Jonas P., Gundelfinger E.D.;
RT "Sequence of D alpha 2, a novel alpha-like subunit of Drosophila
RT nicotinic acetylcholine receptors.";
RL Nucleic Acids Res. 18:3640-3640(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RX MEDLINE=90353591; PubMed=2117557;
RA Jonas P., Baumann A., Merz B., Gundelfinger E.D.;
RT "Structure and developmental expression of the D alpha 2 gene
RT encoding a novel nicotinic acetylcholine receptor protein of
RT Drosophila melanogaster.";
RL FEBS Lett. 269:264-268(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90360975; PubMed=1697262;
RA Sawruk E., Schloss P., Betz H., Schmitt B.;
RT "Heterogeneity of Drosophila nicotinic acetylcholine receptors: SAD,
RT a novel developmentally regulated alpha-subunit.";
RL EMBO J. 9:2671-2677(1990).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Ducan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-X., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Head;
RX MEDLINE=22426066; PubMed=12537569;

RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Krommiller B., Pacle J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
CC Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: CNS in embryos.
CC -!- DEVELOPMENTAL STAGE: Late embryonic and late pupal stages.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X52274; CAA36517.1; -;
CC EMBL; X53583; CAA37652.1; -;
CC EMBL; AE003748; AAF56303.1; -;
CC EMBL; AY058446; AAL13675.1; -;
CC F1R; S11679; ACEFA2.
CC FlyBase; FBgn0000039; nAChR-alpha-96Ab.
CC InterPro; IPR006029; Neu_chan memb.
CC InterPro; IPR006202; Neu_chan LBD.
CC InterPro; IPR006201; Neu_chan.
CC Pfam; PF02931; Neur_chan_LBD; 1.
CC Pfam; PF02932; Neur_chan_memb; 1.
CC PRINTS; PR00252; NR1ONCHANNEL.
CC TIGRFAMs; TIGR00860; LIC; 1.
CC PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
CC Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
CC Transmembrane; Multigene family.
CC SIGNAL 1 21 PROBABLE.
CC CHAIN 22 576 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA-
CC LIKE CHAIN 2.
CC DOMAIN 22 261 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 262 285 POTENTIAL.
CC TRANSMEM 293 311 POTENTIAL.
CC TRANSMEM 327 346 POTENTIAL.
CC DOMAIN 347 526 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 527 545 POTENTIAL.
CC DISULFID 169 183 BY SIMILARITY.
CC DISULFID 243 244 ASSOCIATED WITH RECEPTOR ACTIVATION
CC (BY SIMILARITY).
CC CARBOHYD 65 65 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 254 254 N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 570 570 N-LINKED (GLCNAC...) (POTENTIAL).
CC SEQUENCE 576 AA; 65506 MW; 97D6A46CADCF42F CRC64;
Query Match 35.7%; Score 943; DB 1; Length 576;
Best Local Similarity 36.9%; Pred. No. 2.6e-74;
Matches 197; Conservative 102; Mismatches 189; Indels 46; Gaps 7;
Qy 3 PMLAALALLPVSEOGPEKRLNALLANYNTLRPVAEPEVRFGTLTQIIDV 62
Db 25 PLCLLVLLVLLCETVQANPDAKLYDLLSNVRLRFPVSNNTDTVLVGLGLRLSLQIDL 84
Qy 63 DEKNQLITNWLSEWDSYGVGVKDLRITPNKAWKPDVLMVNSADEGDTG 122
Db 85 NLKQDILITNWLSEWDSYGVGVKDLRITPNKAWKPDVLMVNSADEGDTG 144
Qy 123 YQNVVVRSGGSLYVPPGIFKSTCKMDIAWFFDDQCHDKMFGSWTYDGNQLDLV ---- 178
Db 145 TMTKAILHYTGKVVWTPPAIFKSSCEIDVRYFFPDQQTCTCFMKFGSWTYDGDQIDLKHSQ 204
Qy 179 -----LKDEAGGDLSDFITNGEYVLQWPKGNKITTYACCPPEPVVTFIMIRRLIYY 233


```

Qy 83 NLRWDSYGVKDLRITPNKLMKPDVLMYNSADEGPDGTQTNVVRSGSCLYVPPGI 142
Db 94 KLRWDPMEYDGIETLRVPADKIWKPOILVYNNVAVGDFQVEGKTALLKYNMGMITWTPPAI 153
Qy 143 FKSTCKMDIAWFFPDQCHDMKFGSWTYDGNQDLVLKDEAGG--DLSDFITNGEWYLIIG 200
Db 154 FKSSCPMDITFFPDHQNCSLKFGSWTYDKAEIDLLI--IGSKVDMDNFWENSEWEIID 210
Qy 201 MPKKNITTYACCEPYVDVTFITMIRRTLYYFFNLIVPCVLISSMALLGFTLPDPSGE 260
Db 211 ASGYKHDIKYNCCBEIYTDITYSYIRRLPMFYTNLIIPCLFISFTLVLFVFLPSDCGE 270
Qy 261 KLTGLVTLLSLTVFLNLVAETLPQVSDAIPLLGTYFNCIMFMVYASSVVLTVVVLNYVHR 320
Db 271 KVTLCISVLSLTVFLNLVAETLPQVSDAIPLLGTYFNCIMFMVYASSVVLTVVVLNYVHR 330
Qy 321 TADIHENPOWIKSVFLOLWPLMILMSRPGKKITRKTIMMTRMRELELKERSKSLIANV 380
Db 331 TPTTHTMPRWKTVFLKLLPQVLLMRWPLDK-TRGT--GSDAVPRGLARRPAKGLAS- 385
Qy 381 LDIDDDFRHGGP-----PNNSTASTGNLPGCSIPTDPRRSFVRPS--TMEDVG 428
Db 386 -----HGEPRHLKCFCHKSNELA-----TSKRRLSHQPLQWVVEN-- 422
Qy 429 GGLGSHHREHLHLILRELQFITARMKKADEAEELISDMKFAAMVVDRCFLFVFTLTFTIAT 488
Db 423 ---SEHSPEVEDVINSQVFAENWKSNETKEVEDDMKYVAMVVDVFLWVFLWVFIIVCVFGT 479
Qy 489 VAVLL 493
Db 480 AGLFL 484

```

Search completed: May 7, 2004, 11:36:33
Job time : 10.3565 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:30:03 ; Search time 31.472 Seconds
(without alignments)

5022.709 Million cell updates/sec

Title: US-09-303-232-6

Perfect score: 2640

Sequence: 1 MAPMLAALALLLPVSEQ.....LFTIATVALLSAPHIIQ 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virius:*

16: sp_bacteriapi:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	2640	100.0	501	5	Q9XZ14	Q9xzi4 heliothis v
2	2194	83.1	494	5	Q8T7S1	Q8t7s1 drosophila
3	2181	82.6	494	5	Q8T7S2	Q8t7s2 drosophila
4	2177	82.5	494	5	Q8T7S3	Q8t7s3 drosophila
5	2177	82.5	494	5	Q86MN8	Q86mn8 drosophila
6	2176.5	82.4	509	5	Q8T7S0	Q8t7s0 drosophila
7	2156.5	81.7	523	5	Q8T7R9	Q8t7r9 drosophila
8	1833	69.4	554	5	Q9VL79	Q9vl79 drosophila
9	1803.5	68.3	496	5	Q9XZ13	Q9xzi3 heliothis v
10	1801.5	68.2	542	5	Q86MN7	Q86mn7 drosophila
11	1800.5	68.2	807	5	Q8T7V5	Q8t7v5 drosophila
12	1786.5	67.7	545	5	Q9VW19	Q9vwi9 drosophila
13	1705.5	64.6	525	5	Q81PE2	Q81pe2 drosophila
14	1258.5	47.7	502	11	Q9JHD6	Q9jhd6 mus musculus
15	1254.5	47.5	502	6	Q866A2	Q866a2 macaca mula
16	1245	47.2	555	13	Q7T2U0	Q7t2u0 fugu rubrip

17	1226.5	46.5	511	13	Q03481	Q03481 gallus gall
18	1222	46.3	509	13	Q800C7	Q800c7 brachydanio
19	1214	46.0	513	13	Q7T2R9	Q7t2r9 fugu rubrip
20	1211.5	45.9	486	13	Q7T2S0	Q7t2s0 fugu rubrip
21	1187.5	45.0	554	13	Q7T2T9	Q7t2t9 fugu rubrip
22	1145.5	43.4	480	5	Q81932	Q8i932 caenorhabdi
23	1132	42.9	461	5	Q81932	Q8i932 caenorhabdi
24	1131.5	42.9	474	13	Q7T2U1	Q7t2u1 fugu rubrip
25	1101	41.7	570	5	Q9VJ79	Q9vj79 drosophila
26	1085	41.1	554	5	Q62083	Q62083 caenorhabdi
27	1083.5	41.0	542	5	Q18556	Q18556 caenorhabdi
28	993.5	37.6	505	4	Q86U77	Q86u77 homo sapien
29	985	37.3	335	5	Q9NKO1	Q9nkd1 drosophila
30	982	37.2	537	5	Q8MUR0	Q8mur0 apis mellif
31	978	37.0	515	5	Q46133	Q46133 locusta mig
32	973.5	36.9	537	5	Q9U941	Q9u941 myzus persi
33	970.5	36.8	499	11	Q8R4G9	Q8r4g9 mus musculu
34	970.5	36.8	504	11	Q8BV44	Q8bv44 mus musculu
35	966.5	36.6	499	11	Q8VHH6	Q8vhh6 mus musculu
36	961	36.4	531	5	Q96632	Q96632 heliothis v
37	959.5	36.3	523	5	Q46128	Q46128 heliothis v
38	953.5	36.1	536	5	Q8TOY9	Q8toy9 aplysia cal
39	953	36.1	529	13	Q7T2S4	Q7t2s4 fugu rubrip
40	951.5	36.0	552	5	P91765	P91765 myzus persi
41	951.5	36.0	568	5	Q9NFR5	Q9nfr5 drosophila
42	946.5	35.9	545	5	Q96631	Q96631 heliothis v
43	941	35.6	532	5	Q9U940	Q9u940 myzus persi
44	939.5	35.6	536	5	Q8T9S0	Q8t9s0 aplysia cal
45	939	35.6	533	5	Q8WRS1	Q8wrs1 chilo suppr

ALIGNMENTS

RESULT 1

Q9XZ14

ID Q9XZ14 PRELIMINARY; PRT; 501 AA.

AC Q9XZ14; PRELIMINARY; PRT; 501 AA.

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Putative nicotinic acetylcholine receptor alpha 7-2 subunit.

OS Heliothis virescens (Noctuid moth) (Owlet moth).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;

OC Noctuidae; Heliothinae; Heliothis.

OX NCBI_TaxID=7102;

RN [1]

RP SEQUENCE FROM N.A.

RA Schulte T., Oellers N., Adamczewski M.;

RT "Putative alpha subunits of insect nicotinic acetylcholine receptors

more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than

to other insect nicotinic acetylcholine receptor alpha subunits.";

RL Submitted (APR-1999) to the EMBL/GenBank/DDBI databases.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.

DR EMBL; AF143847; AAD32698.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti...; IEA.

DR GO; GO:0005216; F:ion channel activity; IEA.

DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.

DR GO; GO:0006811; P:ion transport; IEA.

DR GO; GO:0007269; P:synaptic transmission; IEA.

DR InterPro; IPR006201; Neur_chan

DR InterPro; IPR006202; Neur_chan_LBD.

DR Pfam; PF02931; Neur_chan_LBD; 1.

DR Pfam; PF02932; Neur_chan_memb; 1.

DR PRINTS; PR00252; NRIONCHANNEL.

DR TIGRFAMs; TIGR00860; LIC; 1.

DR PROSITE; PS00236; NEUROTRF ION CHANNEL; 1.

KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;

KW Transmembrane.

SQ SEQUENCE 501 AA; 56704 MW; 43CB0DC3960C78AB CRC64;
Query Match 100.0%; Score 2640; DB 5; Length 501;
Best Local Similarity 100.0%; Pred. No. 6.9e-242;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAPMLAALALLPVSEQGPHEKRLNALLANTLERPVANSEPLEVRFGLTLOQII 60
DB 1 MAPMLAALALLPVSEQGPHEKRLNALLANTLERPVANSEPLEVRFGLTLOQII 60
QY 61 DVEKKNQLITNIWLSLEWYNDYLRWDSYGGVKDLRITPNKLMKPDVLMYNSADSGFD 120
DB 61 DVEKKNQLITNIWLSLEWYNDYLRWDSYGGVKDLRITPNKLMKPDVLMYNSADSGFD 120
QY 121 GTYQTNVVRSGGSLYVPPGIFKSTCKMDIAWPFDDQHCMDKFGSWTYDGNOLDIVLK 180
DB 121 GTYQTNVVRSGGSLYVPPGIFKSTCKMDIAWPFDDQHCMDKFGSWTYDGNOLDIVLK 180
QY 181 DEAGGDLSDFTNGEWYLGMPGKNTITVACCEPVVDVTFITMIRRTLYYFFNLIVP 240
DB 181 DEAGGDLSDFTNGEWYLGMPGKNTITVACCEPVVDVTFITMIRRTLYYFFNLIVP 240
QY 241 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYFNCI 300
DB 241 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYFNCI 300
QY 301 MFWVASSVLTVVVLYNHHTADITHMPQWIKSVFLOLWPLWILMSRPGKKITRKTIMN 360
DB 301 MFWVASSVLTVVVLYNHHTADITHMPQWIKSVFLOLWPLWILMSRPGKKITRKTIMN 360
QY 361 TRMELELKERSKSLANVLDDDDFRHGPPPNSTASTGNLPGGCSIFRTDFRSFVR 420
DB 361 TRMELELKERSKSLANVLDDDDFRHGPPPNSTASTGNLPGGCSIFRTDFRSFVR 420
QY 421 PSTMEDVGGGLSHHRELHLILRELQFITARMKKADEAEELISDWKFAAMVVDRCFLVF 480
DB 421 PSTMEDVGGGLSHHRELHLILRELQFITARMKKADEAEELISDWKFAAMVVDRCFLVF 480
QY 481 TLFTIIATVALLSAPHIIQV 501
DB 481 TLFTIIATVALLSAPHIIQV 501

RESULT 2

Q8T7S1 ID Q8T7S1 PRELIMINARY; PRT; 494 AA.
AC Q8T7S1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalph6 subunit variant type III.
DE III.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalph5, Dalph6 and Dalph7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321447; AAM13394.1; -;
DR FlyBase; FBgn0032151; nacr-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.

DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_LBD; 1_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1_memb.
DR Pfam; PF02932; Neur_chan_LBD; 1_memb.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56113 MW; 48327537229573FF CRC64;
Query Match 83.1%; Score 2194; DB 5; Length 494;
Best Local Similarity 82.6%; Pred. No. 1.7e-199;
Matches 418; Conservative 31; Mismatches 37; Indels 20; Gaps 3;
QY 2 APMLAALAL-----LALLPVSEQGPHEKRLNALLANTLERPVANSEPLEVRFGLT 55
DB 3 SPUPASUSLVLLIFLAIIRKESQGPHEKRLNALLANTLERPVANSEPLEVRFGLT 62
QY 56 LQOIIDVDEKNQLITNIWLSLEWYNDYLRWDSYGGVKDLRITPNKLMKPDVLMYNSA 115
DB 63 LQOIIDVDEKNQLITNIWLSLEWYNDYLRWDSYGGVKDLRITPNKLMKPDVLMYNSA 122
QY 116 DEFGDGYQTNVVRSGGSLYVPPGIFKSTCKMDIAWPFDDQHCMDKFGSWTYDGNOL 175
DB 123 DEFGDGYHTNIWVVKHSGSLYVPPGIFKSTCKMDITWFPDDQHCMDKFGSWTYDGNOL 182
QY 176 DLVLKAGGDLSDFTNGEWYLGMPGKNTITVACCEPVVDVTFITMIRRTLYYFF 235
DB 183 DLVNSDGGDLSDFTNGEWYLGMPGKNTITVACCEPVVDVTFITMIRRTLYYFF 242
QY 236 NLIIVPCVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGT 295
DB 243 NLIIVPCVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGT 302
QY 296 YFCIMFWASSVLTVVVLYNHHTADITHMPQWIKSVFLOLWPLWILMSRPGKKITRK 355
DB 303 YFCIMFWASSVLTVVVLYNHHTADITHMPQWIKSVFLOLWPLWILMSRPGKKITRK 362
QY 356 TIMMTRMRELELKERSKSLANVLDDDDFRHGPPPNSTASTGNLPGGCSIFRTDFR 415
DB 363 TILLSNRKMLELKERSKSLANVLDDDDFRH-----TISGQTAIGSS----- 408
QY 416 RSFVRPSTMEDVGGGLSHHRELHLILRELQFITARMKKADEAEELISDWKFAAMVVDRF 475
DB 409 ASFGRTTVEHHTAIGCNHKLHLILRELQFITARMKKADEAEELISDWKFAAMVVDRF 468
QY 476 CLFVFTLTIIATVALLSAPHIIQV 501
DB 469 CLIVFTLTIIATVALLSAPHIIQV 494
RESULT 3
Q8T7S2 ID Q8T7S2 PRELIMINARY; PRT; 494 AA.
AC Q8T7S2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalph6 subunit variant type II.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.

```
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culotto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533 (2002).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321446; AAM13393.1; -.
DR FlyBase; FBgn0032151; nAcR-alpha-30D.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005230; F: extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0005216; F: ion channel activity; IEA.
DR GO; GO:0030594; F: neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P: ion transport; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02932; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56048 MW; 6EE71810E0E7BBB CRC64;

Query Match      82.6%; Score 2181; DB 5; Length 494;
Best Local Similarity 81.6%; Pred. No. 2.9e-198;
Matches 413; Conservative 35; Mismatches 38; Indels 20; Gaps 3;

QY      2 APLMALAL-----LALLPVSEQGPHEKRLNALLANTLERPVANSEPLEVFGLT 55
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      3 SPLPASLSFVLLIFLAIKESQGPHEKRLNHLSTNTLERPVANSEPLEVFGLT 62
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      56 LQOIIDVDEKNQILLTNWLSLEWNDYNLRWNSYGGVKDLRITPNKLPDVLMYNSA 115
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      63 LQOIIDVDEKNQILLTNWLSLEWNDYNLRWNSYGGVKDLRITPNKLPDVLMYNSA 122
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      116 DEGFDTGYTQNVVVRSGGSLYVPPGIFKSTCKMDIAWFPDDQHCDFGWSYDGNQL 175
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      123 DEGFDTGYTQNVVVRSGGSLYVPPGIFKSTCKMDIAWFPDDQHCDFGWSYDGNQL 182
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      176 DLVLKDEAGDLSDFITNGEWLIMGPKKNTIYACCEPVYDVVTFTMIRRTLYYFF 235
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      236 NLIVPCVLISSMALLGFTLPDPSGKLTGLGVITLLSLTVFLNVAETLPQVSDAIPLLGT 295
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      243 NLIVPCVLISSMALLGFTLPDPSGKLTGLGVITLLSLTVFLNVAETLPQVSDAIPLLGT 302
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      296 YFNCIMFVASSVLTIVVNLVYHRTADIHMPQIKSVFLOLWPLRMSPGKKITRK 355
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      303 YFNCIMFVASSVLTIVVNLVYHRTADIHMPQIKSVFLOLWPLRMSPGKKITRK 362
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      356 TIMMTRMRELKERSKSLANVLIDDDDFRHGPPPNSTASTGNLGPSCIFRTDPR 415
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      363 TLLSNRMKELEKERSKSLANVLIDDDDFRH-----TISGQTAIGSS----- 408
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      416 RGFVRSTMEDVGGGLSHHRLHLRLRELQFTIARMKKADEAEELISDWKQAMVVDPRF 475
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      409 ASFGRTTVEEHTAIGCHNKLHLKELQFTIARMKKADEAEELISDWKQAMVVDPRF 468
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      476 CLVFFTLTIIATVAVLLSAPHIIQV 501
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      469 CLIVFTLTIIATVAVLLSAPHIIQV 494
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
Q8T7S3
ID Q8T7S3 PRELIMINARY; PRT; 494 AA.
```

```
AC Q8T7S3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalphas6 subunit variant type I.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culotto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533 (2002).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321445; AAM13392.1; -.
DR FlyBase; FBgn0032151; nAcR-alpha-30D.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005230; F: extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0005216; F: ion channel activity; IEA.
DR GO; GO:0030594; F: neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P: ion transport; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02932; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56095 MW; B46EBEDA63A92942 CRC64;

Query Match      82.5%; Score 2177; DB 5; Length 494;
Best Local Similarity 81.8%; Pred. No. 6.8e-198;
Matches 414; Conservative 33; Mismatches 39; Indels 20; Gaps 3;

QY      2 APLMALAL-----LALLPVSEQGPHEKRLNALLANTLERPVANSEPLEVFGLT 55
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      3 SPLPASLSFVLLIFLAIKESQGPHEKRLNHLSTNTLERPVANSEPLEVFGLT 62
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      56 LQOIIDVDEKNQILLTNWLSLEWNDYNLRWNSYGGVKDLRITPNKLPDVLMYNSA 115
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      63 LQOIIDVDEKNQILLTNWLSLEWNDYNLRWNSYGGVKDLRITPNKLPDVLMYNSA 122
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      116 DEGFDTGYTQNVVVRSGGSLYVPPGIFKSTCKMDIAWFPDDQHCDFGWSYDGNQL 175
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      123 DEGFDTGYTQNVVVRSGGSLYVPPGIFKSTCKMDIAWFPDDQHCDFGWSYDGNQL 182
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      176 DLVLKDEAGDLSDFITNGEWLIMGPKKNTIYACCEPVYDVVTFTMIRRTLYYFF 235
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      183 DLVLNSEDGGLSDFITNGEWLIMGPKKNTIYACCEPVYDVVTFTMIRRTLYYFF 242
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      236 NLIVPCVLISSMALLGFTLPDPSGKLTGLGVITLLSLTVFLNVAETLPQVSDAIPLLGT 295
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      243 NLIVPCVLISSMALLGFTLPDPSGKLTGLGVITLLSLTVFLNVAETLPQVSDAIPLLGT 302
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      296 YFNCIMFVASSVLTIVVNLVYHRTADIHMPQIKSVFLOLWPLRMSPGKKITRK 355
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      303 YFNCIMFVASSVLTIVVNLVYHRTADIHMPQIKSVFLOLWPLRMSPGKKITRK 362
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      356 TIMMTRMRELKERSKSLANVLIDDDDFRHGPPPNSTASTGNLGPSCIFRTDPR 415
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      363 TLLSNRMKELEKERSKSLANVLIDDDDFRH-----TISGQTAIGSS----- 408
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
QY 416 RSFVRPSTMEDVGGGLGSHHREHLILRELQFITARMKKADAEALISDMKFAAMVVDRF 475
Db 409 ASFGRPPTVEHHTAIGCNHKLHLILRELQFITARMKKADAEALISDMKFAAMVVDRF 468
QY 476 CLFVFTFTIIATVAVLLSAPHIIQV 501
Db 469 CLIVFTFTIIATVAVLLSAPHIIQV 494

RESULT 5
Q86MN8 Q86MN8 PRELIMINARY; PRT; 494 AA.
AC Q86MN8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nicotinic acetylcholine receptor subunit Dalpha6 precursor.
GN NACRALPHA-30D.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Millar N.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ554209; CAD86935.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_chan_memb.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
DR SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 494 NICOTINIC ACETYLCHOLINE RECEPTOR SUBUNIT
FT DALPHA6.
FT SEQUENCE 494 AA; 55980 MW; 4DFCS72139587070 CRC64;

Query Match 82.5%; Score 2177; DB 5; Length 494;
Best Local Similarity 81.6%; Pred. No. 6.8e-198;
Matches 413; Conservative 35; Mismatches 38; Indels 20; Gaps 3;

QY 2 APMLAALAL-----LALLPVSEOGPEKRLINALLANYNTLRPVANESPLEVRGLT 55
Db 3 SPLPASUSLVLLIFLAIKESCGPHEKRLINLLSTYNTLRPVANESPLEVRKFLT 62
QY 56 LQIIVDEKNQILITNWLSEWNDYNLNRNDSEYGVKDLRITPNKWKPDVLMVNSA 115
Db 63 LQIIVDEKNQILITNWLSEWNDYNLNRNETEYGVKDLRITPNKWKPDVLMVNSA 122
QY 116 DEGPFGYQTNVVRSGSCLYVPPGIFPKSTCKMDIAWFPDDOHCMDKFGSWYDGNL 175
Db 123 DEGPFGYQTNVVRSGSCLYVPPGIFPKSTCKMDITWFPDDOHCMDKFGSWYDGNL 182
QY 176 DLVLKDAAGDLPDFTNGEWYLMGPKKNTITYACCPPEYVDVTFITMIRRTLYYFF 235
Db 183 DLVLSSDEGGDLPDFTNGEWYLMGPKKNTITYACCPPEYVDITTIQIRRTLYYFF 242
QY 236 NLIIVPCVLISSMALGFTLPDPSGEKLTGLVITLLSLTVFLNVAETLPQVSDAIPLGT 295
Db 243 NLIIVPCVLISSMALGFTLPDPSGEKLTGLVITLLSLTVFLNVAESNFTSDAVPLGT 302
QY 296 YFNCIMFVWASSVLTVVVLYNYHRTADIHMPQWIKSVFLQWLFWILMRSPGKKITRK 355
Db 296 YFNCIMFVWASSVLTVVVLYNYHRTADIHMPQWIKSVFLQWLFWILMRSPGKKITRK 355
```

```
Db 303 YFNCIMFVWASSVLTVVVLYNYHRTADIHMPQWIKSVFLQWLFWILMRSPGKKITRK 362
QY 356 TIMMTRMRELEKERSKSLANVLDDIDDRHGGPPPNSTAGNLGPGCSIFRTDFR 415
Db 363 TILLSNRMLEKERSKSLANVLDDIDDRHGGPPPNSTAGNLGPGCSIFRTDFR 408
QY 416 RSFVRPSTMEDVGGGLGSHHREHLILRELQFITARMKKADAEALISDMKFAAMVVDRF 475
Db 409 ASFGRPPTVEHHTAIGCNHKLHLILRELQFITARMKKADAEALISDMKFAAMVVDRF 468
QY 476 CLFVFTFTIIATVAVLLSAPHIIQV 501
Db 469 CLIVFTFTIIATVAVLLSAPHIIQV 494

RESULT 6
Q8T7S0 Q8T7S0 PRELIMINARY; PRT; 509 AA.
AC Q8T7S0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type IV.
GN NACRALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes
RT Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.
RL Genetics 160:1519-1533(2002).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321448; AA013395.1; -.
DR FlyBase; FBgn032151; nACR-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
DR Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 509 AA; 57987 MW; BE8D8E0198E0C2BD CRC64;

Query Match 82.4%; Score 2176.5; DB 5; Length 509;
Best Local Similarity 80.2%; Pred. No. 8e-198;
Matches 418; Conservative 31; Mismatches 37; Indels 35; Gaps 4;

QY 2 APMLAALAL-----LALLPVSEOGPEKRLINALLANYNTLRPVANESPLEVRGLT 55
Db 3 SPLPASUSLVLLIFLAIKESCGPHEKRLINLLSTYNTLRPVANESPLEVRKFLT 62
QY 56 LQIIVDEKNQILITNWLSEWNDYNLNRNDSEYGVKDLRITPNKWKPDVLMVNSA 100
Db 63 LQIIVDEKNQILITNWLSEWNDYNLNRNETEYGVKDLRITPNKWKPDVLMVNSA 122
QY 101 PNKLWKPDVLMVNSADEGFDGTQYTNVVRSGSCLYVPPGIFPKSTCKMDIAWFPDDQH 160
```

```

Db 123 PNKLKEDVLNYSADSGFGTYHTNIVVRSGLVYVPGIFKSTCKMDITWFPDDQH 182
Qy 161 CDKFGSWTYDGNQLDLVLKDEAGDLSDFITNGEWYILGMPGKNTITYACCPYVDV 220
Db 183 CEMKFGSWTYDGNQLDLVLNSEDGDLSDFITNGEWYLLAMPKNTIVVACCPYVDI 242
Qy 221 TPTIMRRTLYFFNLIYPCVLISSMALLGFTLPDPSGKILGVTIILSLVFLNLVA 280
Db 243 TPTIQRRTLYFFNLIYPCVLISSMALLGFTLPDPSGKILGVTIILSLVFLNLVA 302
Qy 281 ETLFQVSDAIPLLGTYFNCIMFVASSVLTIVVLYNHHRTADITHOMPWKSVFLQWLP 340
Db 303 ETLFQVSDAIPLLGTYFNCIMFVASSVLTIVVLYNHHRTADITHOMPWKSVFLQWLP 362
Qy 341 WILMRSPGKTKITRKTITMNTMRRELKERSKSLIANVLIDDDDFRHPGPPPNSTAST 400
Db 363 WILMRSPGKTKITRKTITMNTMRRELKERSKSLIANVLIDDDDFRHPGPPPNSTAST 415
Qy 401 GNLPGGCSIFRTDPRRSFVRPSTMEDVGGGISHHRELHLILRELQFITARMKKADEAE 460
Db 416 SQTAGSS-----ASFGRTTVEEHHTAIGNCHKDLHLILRELQFITARMKKADEAE 468
Qy 461 LISDWKFAAMVDRFCFLVFTFTIATVAVLLSAPHIIVQ 501
Db 469 LIGDWKFAAMVDRFCFLVFTFTIATVAVLLSAPHIIVQ 509

RESULT 7
Q8T7R9
ID AC Q8T7R9 PRELIMINARY; PRT; 523 AA.
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Nicotinic acetylcholine receptor Dalphae subunit variant type V.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Ephydroidea; Drosophilidae; Drosophila.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culiceto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphae5, Dalphae6 and Dalphae7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533 (2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321449; XAM1396.1; -.
DR FlyBase; FBgn0032151; nACR-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0005216; F:ion channel receptor activity; IEA.
DR GO; GO:000594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTP_ION_CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor; Transmembrane.
SQ SEQUENCE 523 AA; 59110 MW; 1C200AF74F87F841 CRC64;

```

```

Query Match 81.7%; Score 2156.5; DB 5; Length 523;
Best local Similarity 77.6%; Pred. No. 6.6e-196;
Matches 415; Conservative 32; Mismatches 39; Indels 49; Gaps 4;

Qy 2 APLMALAL-----LALLPVSQGHKEKELLNALLANYNTLRPPVANESEPLEVRFGLT 55
Db 3 SPIPASLSLVLLIFALIKESQGHKEKELLNLLSTYNTLRPPVANESEPLEVRFGLT 62
Qy 56 LQOIIDVDEKNQLITNIWLSLEWNDYNLEWNSYGVGKDLRITPNKLPDVLNYSNA 115
Db 63 LQOIIDVDEKNQLITNAWLNLEWNDYNLEWNETEYGVGKDLRITPNKLPDVLNYSNA 122
Qy 116 DEGFDTGYQTNVVRSGSCLVYPPGIFKSTCKMDITWFPDDQCHDKMGSGWTYDGNQL 175
Db 123 DEGFDTGYHTNIVVKGSGCLVYPPGIFKSTCKMDITWFPDDQCHDKMGSGWTYDGNQL 182
Qy 176 DLVLKDEAGDLSDFITNGEWYILGMPGKNTITYACCPYVDVFTIMRRTLYFF 235
Db 183 DLVLNSEDGDLSDFITNGEWYLLAMPKNTIVVACCPYVDITFTIQRRTLYFF 242
Qy 236 NLIVPCVLISSMALLGFTLPDPSGKILT-----GV 266
Db 243 NLIVPCVLISSMALLGFTLPDPSGKILT-----GV 302
Qy 267 TILLSITVFLNLVAETLPQVSDAIPLLGTYFNCIMFVASSVLTIVVLYNHHRTADIHE 326
Db 303 TILLSITVFLNLVAETLPQVSDAIPLLGTYFNCIMFVASSVLTIVVLYNHHRTADIHE 362
Qy 327 MPWIKSVFLOLWPILRMSRPGKTKITRKTITMNTMRRELKERSKSLIANVLIDDD 386
Db 363 MPWIKSVFLOLWPILRMSRPGKTKITRKTITMNTMRRELKERSKSLIANVLIDDD 422
Qy 387 FRHGPPPNSTASTGNLPGGCSIFRTDPRRSFVRPSTMEDVGGGISHHRELHLILRELQ 446
Db 423 FRH-----TISGSAIGSS-----ASFGRTTVEEHHTAIGNCHKDLHLILRELQ 468
Qy 447 FITARMKKADEAEELISDMKFAAMVDRFCFLVFTFTIATVAVLLSAPHIIVQ 501
Db 469 FITARMKKADEAEELISDMKFAAMVDRFCFLVFTFTIATVAVLLSAPHIIVQ 523

RESULT 8
Q9VL79
ID AC Q9VL79 PRELIMINARY; PRT; 554 AA.
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE CG4128 protein.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN SEQUENCE FROM N.A.
RX STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

```


RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lin X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*;"
RL Science 287:2185-2195(2000).
RN [2]
RN SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirska R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome;"
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaninker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.B.;
RT "Annotation of *Drosophila melanogaster* genome;"
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO03626; AAF52817.2; -;
DR FlyBase; FBgn0032151; nAcr-alpha-30D.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_channel.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_channel_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.

DR PROSITE; PS00236; NEUROTR ION_CHANNEL; 1.
SQ SEQUENCE 554 AA; 63236 MW; D6C63EB7F5B32A35 CRC64;
Query Match 69.4%; Score 1833; DB 5; Length 554;
Best Local Similarity 65.0%; Pred. No. 4e-165;
Matches 370; Conservative 36; Mismatches 67; Indels 96; Gaps 8;
QY 4 MLAALALLPVSEOGPEKRLINALLNANTLERPVANESBPFLVRFGLTLQQLIIVD 63
DB 11 LFVLLFLAIKESQCPHEKRLINHLSTYNTLERPVANESBPFLVRFGLTLQQLIIVD 70
QY 64 EKQQLITNLWLEWNYLWVNDSEYGVGKDLRITPNKLPKPDVLMYNSADEGFGTY 123
DB 71 EKQQLITNLWLEWNYLWVNDSEYGVGKDLRITPNKLPKPDVLMYNSADEGFGTY 130
QY 124 QTNVVRSGGCLYVPGIFKSKMDIAWFPDDQHCMDKFGSWTYDGNQLDLVLKDEA 183
DB 131 HTNIVKHNGSCLYVPGIFKSKMDITWFPDDQHCMDKFGSWTYDGNQLDLVLNSED 190
QY 184 GGLSDPITNGEWYLGW-FGKNTITYACCP-----EPYVDVFTTMR 229
DB 191 GGLSDPITNGEWYLGWYHAGKEYSRLRLPTICRYHLYTNSPYI----- 239
QY 230 TLYFFNLIVPCVLISSMALL-----GFTLPDSEKLT----- 263
DB 240 ILFFQFCANCANLIDGPTGLHIAAGFRETDAGRNYTTIINSISKPCRRVHADNVGCCS 299
QY 264 -----LG-----VTLLSLTVLNLVAETLPQVSDAIP 292
DB 300 SYRYTHILKRYSLKYGPIDRTDHYQCRARTVLLSLTVLNLVAETLPQVSDAIP 359
QY 293 LGTYFNCIMFWASSVVLTVVNLNYHRTADIHEMPQWIKSVFLQWLPWILMRGPGKI 352
DB 360 LGTYFNCIMFWASSVVLTVVNLNYHRTADIHEMPQWIKSVFLQWLPWILMRGPGKI 419
QY 353 TRKTMNTMRLELKERSSKLLANVLIDDDFRGPPPNSTAGTGNLGPCCSIFRT 412
DB 420 TRKTILLSNRMKELEKERSKLLANVLIDDDFRH-----TISGQTAIGSS----- 468
QY 413 DFRSFVRPSTMEDVGGGLGSHHRELHLILRLQFITARMKKADEAEELISDWKFAAMV 472
DB 469 ---ASFGRPTTVEHHTTAIGCNHKLHLILKELOFITARMKKADEAEELISDWKFAAMV 525
QY 473 DRFCFLVFTLTIIATVALLSAPHIIIVQ 501
DB 526 DRFCFLVFTLTIIATVALLSAPHIIIVQ 554

RESULT 9

Q9XZ13
ID Q9XZ13 PRELIMINARY; PRT; 496 AA.
AC Q9XZ13;
DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DE 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Putative nicotinic acetylcholine receptor alpha 7-1 subunit.
OS *Heliothis virescens* (Noctuid moth) (Owlet moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Heliothinae; Heliothis.
OX NCBI_TaxID=7102;
RN [1]
RN SEQUENCE FROM N.A.
RA Schulte T., Oellers N., Adamczewski M.;
RT "Putative alpha subunits of insect nicotinic acetylcholine receptors
RT more similar to vertebrate alpha 7 subunits and *C. elegans* Ce21 than
RT to other insect nicotinic acetylcholine receptor alpha subunits.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF143846; AAD32697.1; -;
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.

DR GO; GO:0005216; F:ion channel activity; IEA.
 DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
 DR GO; GO:0006811; P:ion transport; IEA.
 DR GO; GO:0007268; P:synaptic transmission; IEA.
 DR InterPro; IPR006201; Neur_chan.
 DR InterPro; IPR006202; Neur_chan_LBD.
 DR InterPro; IPR006029; Neur_chan_memb.
 DR Pfam; PF02931; Neur_chan_LBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NRTONCHANNEL.
 DR TIGRFAMS; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
 KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
 KW Transmembrane.
 SQ SEQUENCE 496 AA; 56347 MW; 8032FED8515A6210 CRC64;
 Query Match 68.3%; Score 1803.5; DB 5; Length 496;
 Best Local Similarity 68.7%; Pred. No. 2.2e-162;
 Matches 347; Conservative 53; Mismatches 78; Indels 27; Gaps 7;
 QY 2 APLAALLALLPVSEQ-GPHEKRLNALLANYNTLRLPVPANESPLEVRFGLTLQOII 60
 Db 12 APAGLLLLCLLPGRGARGVHEKRLHLLDHYNLERPVVNSDPLQSLFGLTLQOII 71
 QY 61 DVDEKNOLLITNWLSEWYNDYLRWNSYGGVKDLRITPNKLWKPDLVLYNSADEGFD 120
 Db 72 DVDEKNOLLITNWLKLEWDMNLRWNTSDFGVKDLRVPVPHRLWKPDLVLYNSADEGFD 131
 QY 121 GTYQTNVVRSGSCLYVPGIFKSTCKMDIAWFPDDQCHDKMFGSWTYDGNQLDLVK 180
 Db 132 STYPTNVVRNNGSCLYVPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGYQLDLQ 191
 QY 181 DEAGGLSDFITNGEWYLGMPGKNTITYACPEPVVDVTFITMIRRTLYYFENLIVP 240
 Db 192 DEGGIISFVTNGEWYLGMPGKNEIYVNCPEPYDITFAVIRKRTLYYFENLIVP 251
 QY 241 CVLISSMALLGFTLPDPSGKLTGLVTLLSLVFLNVAETLPQVSDAIPLLGTGFNCI 300
 Db 252 CVLISSMALLGFTLPDPSGKLSGLVTLLSLVFLNVAETLPATSDAVPLIGTGFNCI 311
 QY 301 MFVAVSVLVTVVLYNHHRTADIHEMPQWIKSVFLQWLFWILRMGRPGKKITRKTIMN 360
 Db 312 MFVAVSVVSTIILYNYHRRHADTHEMSDWIRCVFLYWLFWILRMGRPGSATTPPARV- 370
 QY 361 TRMRELKERSKSLANVLDDDDFRHGGPPPNSTASTGNLGPSCIFRTDFRSFVR 420
 Db 371 PPPDLELERSKSLANVLDDDDFRH-----PQAOQ-----PQCRY---YRGG--- 414
 QY 421 PSTMEDVGGGLGSH-----HRELHLILRELQFITARMKKADEAEALISDWKFAAMVVDRF 475
 Db 415 -----EENGAGLAHSCFGVDYELSLILKEIRVITDQWRKDEADADISRDWKFAMVVDEL 470
 QY 476 CLFVFTLTIIATVAVLLSAPHIV 500
 Db 471 CLIIFTLTIIATVAVLLSAPHIMV 495

RESULT 10
 Q86MN7 PRELIMINARY; PRT; 542 AA.
 ID Q86MN7
 AC Q86MN7
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Nicotinic acetylcholine receptor subunit Dalpha7 precursor.
 GN NACRALPHA-18C.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Millar N.S.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ554210; CAD86936.1; -.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
 DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
 DR GO; GO:0006811; P:ion transport; IEA.
 DR InterPro; IPR006201; Neur_chan.
 DR InterPro; IPR006202; Neur_chan_LBD.
 DR InterPro; IPR006029; Neur_chan_memb.
 DR Pfam; PF02931; Neur_chan_LBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NRTONCHANNEL.
 DR TIGRFAMS; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
 KW Signal; Receptor.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 542 NICOTINIC ACETYLCHOLINE RECEPTOR SUBUNIT
 FT DALPHA7.
 SQ SEQUENCE 542 AA; 60988 MW; 69D2C39746BB74D7 CRC64;
 Query Match 68.2%; Score 1801.5; DB 5; Length 542;
 Best Local Similarity 65.5%; Pred. No. 3.9e-162;
 Matches 347; Conservative 59; Mismatches 79; Indels 45; Gaps 5;
 QY 4 MLAALALLPVSEQGPHEKRLNALLANYNTLRLPVPANESPLEVRFGLTLQOIIDVD 63
 Db 20 LVYGLGLLIMIPACAAAPHEKRLHLLDNYSLRLPVPANESDPLQSLFGLTLQOIIDVD 79
 QY 64 EKNOLLITNWLSEWYNDYLRWNSYGGVKDLRITPNKLWKPDLVLYNSADEGPDGT 123
 Db 80 EKNOLLITNWLKLEWDMNLRWNSDFGVGRDLRIPPHRLWKPDLVLYNSADEGPDGT 139
 QY 124 QTNVVRSGSCLYVPGIFKSTCKMDIAWFPDDQCHDKMFGSWTYDGNQLDLVKDEA 183
 Db 140 ATNVVRNNGSCLYVPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGFQDLQLDEA 199
 QY 184 GGLSDFITNGEWYLGMPGKNTITYACPEPVVDVTFITMIRRTLYYFENLIVPCVL 243
 Db 200 GDISFITNGEWDLGVGKNEIYVNCPEPYDITFAILIRKRTLYYFENLIVPCVL 259
 QY 244 ISSMALLGFTLPDPSGKLTGLVTLLSLVFLNVAETLPQVSDAIPLLGTGFNCIMFM 303
 Db 260 IASMALLGFTLPDPSGKLSGLVTLLSLVFLNVAETMPATSDAVPLIGTGFNCIMFM 319
 QY 304 VASSVVLTVVLYNHHRTADIHEMPQWIKSVFLQWLFWILRMGRPGK-----KI 352
 Db 320 VASSVVTIILYNYHRRPDTHEMSDWIRVFLYWLFWILRMGRPGQVGECEPPSSSS 379
 QY 353 TRKTIMNTRMRELKERSKSLANVLDDDDFRHGGPPPNSTASTGNLGPSCIFRT 412
 Db 380 SSASGEKKQOIQNVLEKERSKSLANVLDDDDFR-----CNHRCASATLPHQTYRT 434
 QY 413 DFRRSFVRPSTMED-----VGGGL-----GSHHRELHLILRELQFITA 450
 Db 435 MYRQG-----DDGSVGVGPAGPVVDGRLHEAHSHTCLTSSAEYELALILKEUWITE 487
 QY 451 RMKKADEAEALISDWKFAAMVVDRFCLFVFTLTIIATVAVLLSAPHIV 500
 Db 488 QLKXDETDITRDWKFAAMVVDRCLIIFTLTIIATVAVLSAPHIV 537

RESULT 11
 Q8T7V5 PRELIMINARY; PRT; 807 AA.
 ID Q8T7V5
 AC Q8T7V5
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Nicotinic acetylcholine receptor Dalphas subunit.
 GN NACR-ALPHA-34E OR NACRALPHA-34E OR BG:DS05899.4 OR BG:DS05899.5 OR
 GN CG4498 OR CG16878 OR CG32975.
 OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

RN [1]

RN SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalpaa5, Dalpaa6 and Dalpaa7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genomics 160.1519-1533 (2002).

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.

CC ENBL; AF272778; AAM13390.1; -.

DR FlyBase; FBgn028875; nAChR-alpha-34E.

DR GO; GO:0016021; C: integral to membrane; IEA.

DR GO; GO:0005230; F: extracellular ligand-gated ion channel activity; IEA.

DR GO; GO:0005216; F: ion channel activity; IEA.

DR GO; GO:0030594; F: neurotransmitter receptor activity; IEA.

DR GO; GO:0006811; P: ion transport; IEA.

DR GO; GO:0007268; P: synaptic transmission; IEA.

DR InterPro; IPR006201; Neur_chan_LBD.

DR InterPro; IPR006029; Neur_chan_LBD.

DR Pfam; PF02931; Neur_chan_LBD; 1.

DR Pfam; PF02932; Neur_chan_memb; 1.

DR PRINTS; PR00252; NRIONCHANNEL.

DR TIGRFAMs; TIGR00860; ITC; 1.

DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.

DR GlycoProtein; GlycoProtein; Post-synaptic membrane; Receptor;

DR Transmembrane. Ionic channel; Post-synaptic membrane; Receptor;

DR SEQUENCE 807 AA; 91223 MW; CBB4F6B34287C8C8 CRC64;

Query Match 68.2%; Score 1800.5; DB 5; Length 807;

Best Local Similarity 66.3%; Pred. No. 8.6e-162;

Matches 348; Conservative 61; Mismatches 69; Indels 47; Gaps 9;

10 LLALLPVSEQ-----GPHKKLLMALLANYTLERPVANESPELVREGLTLQOIIDVDEK 65

295 LLIYLNLSAKVCLAGYHEKRLHLLDPYNTLPRVLSNPLQSLFGLTLQOIIDVDEK 354

66 NQLITNWLISLENDYLNWDSYGVKDLRTIPNKLKDPDVLNMSADEGDFGTQY 125

355 NQLIVTNWLENDYLNWDSYGVKDLRTIPNKLKDPDVLNMSADEGDFGTQY 414

126 NVVVRSGSCLYVPPGIFKSTCKMDIAWFPDDQCDMKFGSWTYDGNQLDLVLKDBAGG 185

415 NVVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQCDMKFGSWTYDGNQLDLVLKDBAGG 474

186 DLSDFTNGEWLYLGMGKNTIYACCPPEYVDVTFIMTRRTLYFFNLIYPCVLIS 245

475 DISSVYLNGLWELLVGPKRNEIYNCPPPEYIDITFAIIRRTLYFFNLIYPCVLIA 534

246 SMALLGFTLPDSEKSLGVTLLSLTVFLNVAETLPQVSDAIPILGTGFNCIMFWA 305

535 SMALLGFTLPDSEKSLGVTLLSLTVFLNVAETLPQVSDAIPILGTGFNCIMFWA 594

306 SSVVLTVVVLYNHRHTADIHEMPONIKSVFQLPWLIRMSRPGKKITRK---TIMMNR 362

595 SSVVSTILLNHYHNRADTHMSEWIRIVFLCWLPLWIRMSRPGKRIEPTTPCSDTS 654

363 -----MRELEKERSKSLANVLDDDDFRHG---PPPNSTAGTNGLPGGSIPTD 413

655 SERKHQILSDVELKERSKSLANVLDDDDFRHNCRPWTG-----GLTHPNFAFYIT 709

414 FRRSFVRSTMEDVG---GGIGS-----HHR-----ELHLILRELOFITARMKKA 455

710 YGQ-----DDGSIIGISGRTPMDAVTHHTCIKSTSEYELGLILKEIRFITDLQRKD 761

456 DEEALISDNKFAAMVVRFCFLVFVTLFTIATVAVLLSAPHIIV 500

762 DECNDIANDWKFAAMVVRCLLIFTWFAILLVLLSAPHIIV 806

RESULT 12

Q9VWI9

ID Q9VWI9

PRELIMINARY;

PRT; 545 AA.

AC Q9VWI9;

DT 01-MAY-2000 (T-EMBLrel. 13, Created)

DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)

DE CG32538 protein.

GN NCAR-ALPHA-18C OR CG8082 OR CG8109 OR CG32538.

OS Drosophila melanogaster (fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OC NCBI_TaxID=7227;

RN [1]

SEQUENCE FROM N.A.

STRAIN=Berkeley;

MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Balow R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferraz C., Fleischmann W.,

RA Fodor C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali B., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Klamel M.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RT Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195 (2000).

RN [2]

SEQUENCE FROM N.A.

RA Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,

RA Evans C.A., Gocayne J.D., Anantides P.G., Brandon R.C., Rogers Y.,

RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,

RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,

RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,

RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,

RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,

RA Ibegwam C., Jalali M., Kruse D., Li P., Matti B., Moshrefi A.,

RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,

RA Pacle J., Paragas V., Park S., Patel S., Pfeiffer B.,

RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,

RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,

Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
"Sequencing of Drosophila melanogaster genome";
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kamnik J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Bergman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RA "Annotation of Drosophila melanogaster genome";
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A.
RA FlyBase;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RA EMBL; AE003311; AAF48950.2; nAcra-alpha-18C.
RA FlyBase; FBgn0031014; nAcra-alpha-18C.
RA GO; GO:0016020; C:membrane; IEA.
RA GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
RA GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
RA GO; GO:0006811; P:ion transport; IEA.
RA InterPro; IPR006201; Neur_channel.
RA InterPro; IPR006202; Neur_chan_LBD.
RA InterPro; IPR006029; Neur_channel_memb.
RA Pfam; PF02931; Neur_chan_LBD; 1.
RA Pfam; PF02932; Neur_chan_memb; 1.
RA PRINTS; PR00252; NRIONCHANNEL.
RA TIGRFS; TIGR00860; LIC; 1.
RA PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
RA SEQUENCE 545 AA; 61517 MW; 7B83025107A66209 CRC64;
Query Match 67.7%; Score 1786.5; DB 5; Length 545;
Best Local Similarity 66.7%; Pred. No. 1e-160;
Matches 343; Conservative 56; Mismatches 70; Indels 45; Gaps 5;
20 GPHEKLLNALLANYTLERPVANSEPLEVREGLTQIIVDDEKXQLITNWLISLEW 79
39 GPHEKLLHALLDNYSLSRPVNSDPLQSLGTLMIQIIVDDEKXQLITNWLKLEW 98
80 NDYNLRWNSSEYGVKDLITPNKLPDVLNNSADEGPDGYQYNVVRSGSCLYVP 139
99 NDNLRWNSSEYGVKDLITPNKLPDVLNNSADEGPDGYQYNVVRSGSCLYVP 158
140 PGIFKSTKMDJAWFPDQDCMKFGSWTYDGNQLDLVLKOBAGDLSDFITNGEWYLI 199
159 PGIFKSTKIDITWFFDQRCMKFGSWTYDGFQDLQDLQDAGGDISFITNGEWYLI 218
200 GMPGKKNITTYACCPYVDVITIMRRRTLYFFNLVPCVLISSMALLGFTLPDGS 259
219 GYEGKRENYVNCPEYIDITFALLIRKRLYFFNLVPCVLISMAALLGFTLPDGS 278
260 EKLTLGVITLLSLTVPNLVAETLPVDSDAIPLLGYFNCIMFWASSVVLTVVLYNHH 319
279 EKLTLGVITLLSLTVPNLVAETLPVDSDAIPLLGYFNCIMFWASSVVLTVVLYNHH 338
320 RTADIHEMPQWTKSVFLOLWPLMRSPGK-----KIFKTIIMTRREL 368
339 RNPDTHEMSEWIRVFLYLPCLLRMQRGVQVCEPFPSSSSSSASGEKQIQNVEL 398
369 KERSKSLIANVLIDDDPHGPPPNSTASTGNLPGCSIFRTPRRFRVAPSTMED-- 426
399 KERSKSLIANVLIDDDPHGPPPNSTASTGNLPGCSIFRTPRRFRVAPSTMED-- 446
427 -----VGGL-----GSHREHLILRELQFTARMKKADEAEIISDWK 466
447 VGFVGPAGVPVGRHLHEAHSHTCLTSSAEYELALILKELRWITEQLKKEDETSITRDWK 506

Qy 467 FAAMVDRPCLFVFTFTIATVALLSAPHIIV 500
Db 507 FAAMVDRCLIIFTFTIATVALLSAPHFIV 540

RESULT 13

Q8IPE2 PRELIMINARY; PRT; 525 AA.
AC Q8IPE2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG4128-PC.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dudson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL SEQUENCE 287:2185-2195(2000).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Dou P.L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez C., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacle J., Paragas V., Park S., Patel S., Pfeiffer B.,


```

Db 130 VLVNASGHCQYLPPIGFIKSSCYIDVRWFFEDVQCKLKFGSKSYGWSLQLQMQE--AD 186
Qy 187 LSDFITNGEWYLIQMPGKKNITTYACCPPEYVDVFTTIRRRLLYYFENLIVPCVLISS 246
Db 187 ISSIPNGEWLDMGIPGRKNEKFYECCKEPYDPVTYTTVMRRRTLYGILNLLIPCWLISA 246
Qy 247 MALLGFTLPPDSGKLTGVTLLISLTVFLNLVAETLPQVSDAIPLLGTYFNCIMFMVAS 306
Db 247 LALLVFLPADSGEKISIGITVLLSLTVFLLVAEIMPATSDSVPLIAQYFASMTIIVGL 306
Qy 307 SVLVTVVVLNVHHTADIHMPQWKSVFLOLWFLMRLMSPGKKITRKTIMMTRMREL 366
Db 307 SVVTVIVLVYHHHDPDGGKMPKWTIRILLNWCWFLRMKRPGEKVRPACQHKPRCSL 366
Qy 367 ELKERS--KSLANVLDDDDFR-----HGPPPNSTASTGNLPGGCSIFRTDFRSFV 417
Db 367 ASVELSAGAPPSNGNLLYI--GFRLEGMEHCAPTDPDSGVVCGRL--ACSPTHDEHLMH 422
Qy 418 FVRPSTMEDVGGGLGSHHRELHLILRELQFTTARMKXADEAEELISDWKFAAMVVD3FCL 477
Db 423 GTHPSDGP-----DLAKILEVRYIANFRCDSEVICSEWKFACVVDRLCL 472
Qy 478 FVFTLFTIATVALLSAPHII 499
Db 473 MAFSVFTIICITIGILMSAPNFV 494

```

RESULT 15

```

Q866A2 PRELIMINARY; PRT; 502 AA.
AC Q866A2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nicotinic acetylcholine receptor alpha7.
GN CHRNA7.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Proskocil B.J., Sekhon H.S., Keller J.A., Jia Y., Blakely R.D.,
RA Lindstrom J., Spindel E.R.;
RT "An Intrinsic Non-neuronal Nicotinic Cholinergic Signaling System in
RL Monkey Lung Airway Epithelium.";
DR EMBL; AF486623; AAC084497.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0008811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur channel.
DR InterPro; IPR006202; Neur chan LBD.
DR InterPro; IPR006029; Neu Channel LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Receptor.
SQ SEQUENCE 502 AA; 56429 MW; 213C8A28224AC4A CRC64;

```

```

Query Match 47.5%; Score 1254.5; DB 6; Length 502;
Best Local Similarity 49.6%; Pred. No. 3.4e-110;
Matches 249; Conservative 79; Mismatches 147; Indels 27; Gaps 7;
Qy 8 LALLA-LLPVSQGPHEKELLNALLANYTLRPPVANSEPLEVRPGLTLOQIIVDEKN 66
Db 10 LALAASLLHVSLSQGFQRLKYLKVNPNPLRPVANDSQPLTVYFSLQLQIMDVDEKN 69

```

```

Qy 67 QLLITNIWLSLEWNDYNIWRNDSEYGGVKDILRITPNKLWPKDVLMYNSADEGFDGTQYTN 126
Db 70 QVLTATNIWLOMSWTDHYLQMNVSSEYGVKTVRFPDGQIWKPDILLYNSADERFDATHTN 129
Qy 127 VVRSRSGSCLVBPICFKSTCKMDIANWFFDDQCHDKMFGSWTYDGNOLDLVLKDEAGD 186
Db 130 VLVNSSHCHQYLPPIGFIKSSCYIDVRWFFEDVQCKLKFGSKSYGWSLQLQMQE--AD 186
Qy 187 LSDFITNGEWYLIQMPGKKNITTYACCPPEYVDVFTTIRRRLLYYFENLIVPCVLISS 246
Db 187 ISSIPNGEWLDMGIPGRKNEKFYECCKEPYDPVTYTTVMRRRTLYGILNLLIPCWLISA 246
Qy 247 MALLGFTLPPDSGKLTGVTLLISLTVFLNLVAETLPQVSDAIPLLGTYFNCIMFMVAS 306
Db 247 LALLVFLPADSGEKISIGITVLLSLTVFLLVAEIMPATSDSVPLIAQYFASMTIIVGL 306
Qy 307 SVLVTVVVLNVHHTADIHMPQWKSVFLOLWFLMRLMSPGKKITRKTIMMTRMREL 366
Db 307 SVVTVIVLVYHHHDPDGGKMPKWTIRILLNWCWFLRMKRPGEKVRPACQHKPRCSL 366
Qy 367 ELKERS--KSLANVLDDDDFR-----HGPPPNSTASTGNLPGGCSIFRTDFRSFV 419
Db 367 ASVELSAGAPPSNGNLLYI--GFRLEGMEHCAPTDPDSGVVCGRL--ACSPTHDEHLMH 422
Qy 418 FVRPSTMEDVGGGLGSHHRELHLILRELQFTTARMKXADEAEELISDWKFAAMVVD3FCL 477
Db 414 -PTHDEHLLHGQPPGEGDPLAKILEVRYIANFRCDSEAVCSWKFACVVDRLCL 472
Qy 478 FVFTLFTIATVALLSAPHII 499
Db 473 MAFSVFTIICITIGILMSAPNFV 494

```

Search completed: May 7, 2004, 11:38:39
Job time : 32.472 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 11:50:34 ; Search time 3498.21 Seconds
(without alignments)
4276.744 Million cell updates/sec

Title: US-09-303-232-6
Perfect score: 2640
Sequence: 1 MAPMLAALALLLPVSEQ.....LFTIATVAVLLSAPHIIQ 501

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cpn2_1/USPTO_spool/US09303232/runat_07052004_101110_23893/app_query.fasta_1.2261
-DB=EST -QPMI=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303232 @CGN 1.1 4577 @runat_07052004_101110_23893 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estlin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gas_hum.*
18: em_gas_inv.*
19: em_gas_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C					
1	1273	48.2	885	12	BG632919 GH16126.3
2	970.5	36.8	1864	11	AK053497 Mus muscu
3	970.5	36.8	2916	11	AK051730 Mus muscu
4	970.5	36.8	3126	11	AK080415 Mus muscu
5	957.5	36.3	1436	29	AY402873 Homo sapi
6	944.5	35.8	1436	29	AY402875 Mus muscu
7	924	35.0	2010	11	AK080475 Mus muscu
8	922.5	34.9	2296	14	CD013901 90134548
9	914	34.6	1454	29	AY402878 Mus muscu
10	910.5	34.5	4037	11	AK049722 Mus muscu
11	910.5	34.5	4046	11	AK051742 Mus muscu
12	909	34.4	1201	9	AL530299 AL530299
13	905	34.3	1374	29	AY406230 Homo sapi
14	890.5	33.7	2940	11	AK034228 Mus muscu
15	890.5	33.7	3230	11	AK083157 Mus muscu
16	886.5	33.6	4290	11	AK029177 Mus muscu
17	886	33.6	1374	29	AY406232 Mus muscu
18	884	33.5	1442	29	AY402876 Mus muscu
19	882.5	33.4	1781	14	CD013889 90138814
20	881	33.4	1442	29	AY402877 Pan trogl
21	865.5	32.8	1374	29	AY406231 Pan trogl
22	853	32.3	4589	11	AK030464 Mus muscu
23	849	32.2	922	13	BU915857 AGENCOURT
24	847	32.1	607	9	AI292581 GH15518.5
25	846.5	32.1	1436	29	AY402874 Pan trogl
26	838.5	31.8	3483	11	AK081254 Mus muscu
27	791.5	30.0	1466	14	CD013891 90139069
28	776.5	29.4	2257	11	AK017571 Mus muscu
29	770	29.2	1751	14	CD013905 90134608
30	770	29.2	2110	14	CD013902 90134640
31	770	29.2	2110	14	CD013903 90134855
32	770	29.2	2513	11	AK033068 Mus muscu
33	769.5	29.1	2074	11	AK087554 Mus muscu
34	759	28.7	908	13	BU149265 AGENCOURT
35	748.5	28.4	1034	13	BU149265 AGENCOURT
36	737.5	27.9	1596	14	CD013888 BX403124 BX403124
37	716	27.1	1603	29	AY411327 Homo sapi
38	703.5	26.6	833	14	CB245337 UI-M-FYO
39	693	26.2	902	29	AY407184 Homo sapi
40	691.5	26.2	2948	11	AK081344 Mus muscu
41	686	26.0	902	29	AY407186 Mus muscu
42	682.5	25.9	615	14	CB149460 K-EST0205
43	679.5	25.7	1603	29	AY411329 Mus muscu
44	672.5	25.5	1835	11	AK010496 Mus muscu
45	666	25.2	1556	14	CD013895 90119754

ALIGNMENTS

RESULT 1
BG632919/c
LOCUS BG632919
DEFINITION GH16126.3prime GH Drosophila melanogaster head poly2 Drosophila melanogaster cDNA clone GH16126.3 similar to CG4128: FBand004128 'ion channel', located on: 2L 30D1-30E1:: 04/10/2001, mRNA sequence.
ACCESSION BG632919
VERSION BG632919.1 GI:13758409
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephyrdoidea; Drosophilidae; Drosophila.
1 (bases 1 to 885)
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Teang, G.,
Lewis, S. and Rubin, G.M.
BDGP/HMMI Drosophila EST Project
Unpublished (2001)
Other ESTs: GH16126.5prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd. Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Based upon the presence of a xhoI site followed by a run of 14 or
more T residues at the beginning of the sequence, this clone was
polyadenylated. The resulting poly-T sequence has been removed. hit
genomic AE003511: arm.X [18792641..19136447]
estimated-cyto.18A3-18C6: 04/10/2001
Plate: GH.161 row: C column: 2
High quality sequence stop: 784.
Location/Qualifiers
source
1..885
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clones="GH16126"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/clone_lib="GH Drosophila melanogaster head pOT2"
/note="Organ: head; Vector: pOT2; Site 1: EcoRI; Site 2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2. Plasmid cDNA library."

Alignment Scores:
Pred. No.: 5,05e-139 Length: 885
Score: 1273.00 Matches: 227
Percent Similarity: 91.01% Conservative: 26
Best Local Similarity: 81.65% Mismatches: 25
Query Match: 48.22% Indels: 0
Gaps: 12

US-09-303-232-6 (1-501) x BG632919 (1-885)

Qy 54 LeuThrLeuGlnGlnIleAspValAspGluLysAsnGlnLeuLeuIleThrAsnIle 73
Db 885 CTAACACTCATGCAGATTATCGATGTGCACGAGAGAAATCAACTGCTTATAACGAATATT 826

Qy 74 TrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTyrAsnAspSerGluTyrGly 93
Db 825 TGGCTCAAATTTGGAATGGAACGATATGAATCTTCGATGGAATTCGAGTGAGTTCGGTGT 766

Qy 94 VallyAspLeuArgIleThrProAsnLysLeuTrpLysProAspValLeuMetTyrAsn 113
Db 765 GTGGGGATCTGGGNAATTCGGCCATCGCTATCGAAACCGAGTACTGATGATACAC 706

Qy 114 SerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValArgSerGlyGly 133
Db 705 AGTGGCCGACGAGGGCTTCGATGGAACGATGACGACCAAAATGTGGTTCGCAATAATGGG 646

Qy 134 SerCysLeuTyrValProGlyIlePheLysSerThrCysLysMetAspIleAlaTrp 153
Db 645 AGCTGTCTGTACGTACCGCCAGGTATATTAACTCACTGTAAGTCACTGTAAGTCACTG 586

Qy 154 PheProPheAspAspGlnHisCysAspMetLysPheGlySerTyrThrTyrAspGlyAsn 173
Db 585 TTTCATTTCGACGATCAGAGATGTGAATGAAATTTGGTTCGTGGACCTACCATGGTTT 526

Qy 174 GlnLeuAspLeuValLeuLysAspGluAlaGlyGlyAspLeuSerAspPheIleThrAsn 193
Db 525 CAGTTGGACCTGCAGTTGCAGGACGAAGCTGGTGGCAGCAATTTCTAGCTTTATAACCAAT 466

Qy 194 GlyGluTyrTyrLeuIleGlyMetProGlyLysLeuAsnThrIleThrTyrAlaCysCys 213
Db 465 GCGGAATGGGACTTGTAGTGTGCCGGTAAACGAATGAAATCTACTATAATTTGCTGC 406

Qy 214 ProGluProTyrValAspValThrPheThrIleMetIleArgArgThrLeuTyrTyr 233
Db 405 CCAGAACTTATATTGACATAACATTTCGCAATTTTGTAGGGCGCAAAACGTTGTACTAT 346

Qy 234 PhePheAsnLeuIleValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThr 253
Db 345 TTTTTCATCTGATTGTGGCGTGGTACTATCGCTCCATGGCACTGCTAGGGTTTACA 286

Qy 254 LeuProProAspSerGlyGluLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThr 273
Db 285 CTGCCACCATTTCTGTGAAAGCTTCGCTTGGAGTTTACAATTTCTATTATCGCTTACA 226

Qy 274 ValPheLeuAsnLeuValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeu 293
Db 225 GTCTTCTCAACATGGTGGCGGAAACAATCCGCGACCTCCGATCGGTACCGCTGCTC 166

Qy 294 GlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValValLeuThrValVal 313
Db 165 GGAACCTATTTCATTCGATATGTTATGTTGGCTCATCAGTTGTGTCAACCATATT 106

Qy 314 ValLeuAsnTyrHisArgThrAlaAspIleHisGluMetProGlnTrpIle 331
Db 105 GTCTCAATTCATCATAGAAATCCAGATACGCAATGAAATGAGTGAATGGTA 52

RESULT 2
AK053497 1864 bp mRNA linear HTC 20-SEP-2003
AK053497
LOCUS
DEFINITION
Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
library, clone: E130103E14 product: NEURONAL NICOTINIC ACETYLCHOLINE
RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.

ACCESSION
VERSION AK053497.1 GI:26343494
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159
REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Warahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED
11076861
REFERENCE
4 The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
JOURNAL

[illegible]

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Kikazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

4 The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

⁵ The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
⁶ (bases 1 to 3126)

Db 968 ACGTCTGCTATCGTCTCTCTCCCTGAGCGGCTTTCTCTCGTGATCACCGAGACC 1027
QY 283 LeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPhe 302
Db 1028 ATCCCTCCACTCAGTGTATCCCTTCGATCGGGAGTACCTCTCTTCACTATGATT 1087
QY 303 MetValAlaSerSerValValLeuThrValValValLeuAsnTyrHisHisArgThrAla 322
Db 1088 TTGTGACCTGTTCATCGTATCACAGCTTTGTGCTCAACGTCACACAGAACTCCG 1147
QY 323 AspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrpLeuProTrpIle 342
Db 1148 ACCACACACAGATGCCACTTGGGTCAAGCTGTGTTTGAACCTTCTCCCGAGGTC 1207
QY 343 LeuArgMetSerArgProGlyLysLysIleThrArgLysThrIleMetMetAsnThrArg 362
Db 1208 ATGTTATGATAGTACGACCAACGACCGAGGAAGACGCCGCCAAGAGAGAACTTCTAC 1267
QY 363 MetArgGluLeuGlu-----LeuLysGluArgSerSerLysSerLeuLeuAla 378
Db 1268 GGTGCTGAGCTCTCAACCTGAACCTGCTTCAGCGTCGACAGTCTCAAAAGCTGC----- 1321
QY 379 AsnValLeuAspIleAspAspAspPheArgHisGlyProProProProAsnSerThrAla 398
Db 1322 -----AAGGAAGGCTACCCCTGCCAAGATGGGACATGT 1354
QY 399 SerThrGlyAsnLeuGlyProGlyCysSerIlePheArg-----ThrAspPheArg 415
Db 1355 GGCTAT-----TGCACACCGTAGGGTAAATAATCTCAAAATTTTCAGT 1396
QY 416 ArgSerPheValArgProSerThrMetGluAspValGlyGly-----GlyLeuGlySer 433
Db 1397 GCCAACTCTCAAGAAGCTCCAGTCTGAGTCTGTTGATCTGTGTGCTGCTGCTGCT 1456
QY 434 HisHisArgGluLeuHisLeuLeuLeuArgGluLeuGlnPheIleThrAlaArgMetLys 453
Db 1457 CTGTACACAGAAATCAAGAAGCATCCAAAGTGTGAATATATTGCTGAATAATGAAA 1516
QY 454 LysAlaAspGluGluAlaGluLeuIleSerAspTrpLysPheAlaAlaMetValValAsp 473
Db 1517 GCACAGATGTAGCAAGAGATTCAGATGATTTGGAAGTATGTTGCCATGGTGAATGAT 1576
QY 474 ArgPheCysLeuPheValPheThrLeuPheThrIleIleAlaThrValAlaValLeuLeu 493
Db 1577 CGCATTTTCTCTGGGTTTTCATCCTGTGTGCTGATTTAGCAATGCAGGATTAATCTG 1636

RESULT 5
LOCUS AY402873 1436 bp DNA linear GSS 15-DEC-2003
DEFINITION Homo sapiens CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY402873
VERSION AY402873.1 GI:39759856
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1436)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
TITLE

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..1436
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1436
/gene="CHRNA3"
/locus_tag="HCM1369"
gene
ORIGIN
Alignment Scores:
Pred. No.: 2,2e-101 Length: 1436
Score: 957,50 Matches: 199
Percent Similarity: 61,55% Conservative: 94
Best Local Similarity: 41,81% Mismatches: 165
Query Match: 36,27% Indels: 21
DB: 29 Gaps: 6

US-09-303-232-6 (1-501) x AY402873 (1-1436)

QY 23 GluLysArgLeuLeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAla 42
Db 21 GAGCACGCTATTTCAGCGCTGTTTGAAGATTACAATGAGATCATCCGCGCTGTGGCC 80
QY 43 AsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleLeuAspVal 62
Db 81 AAGTGTCTGACCCAGTCATCCATTCGAGGTGCCATGTCCTAGCTGGTGAAGGTG 140
QY 63 AspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsnAspTyr 82
Db 141 GATGAAGTAAACCCAGATCATCGAGACCACTGTGCTCAAGCAATCTGGAATGACTAC 200
QY 83 AsnLeuArgTrpAsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsn 102
Db 201 AAGCTGAAGTGAACCCCTCTGACTATGTTGGGCGAGTTCATCGCTGCCGTGCACAG 260
QY 103 LysLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThr 122
Db 261 AAGATCTGGAGCCACACATTTGTGCTGATAAACAATGCTTTGGGGATTTCCAGGTGGAC 320
QY 123 TyrGlnThrAsnValValValArgSerGlyGlySerCysLeuTyrValProGlyIle 142
Db 321 GACAAGACCAAGCCCTTACTCAAGTACACTGGGAGGTGACTTGGATACCTCCGGCCATC 380
QY 143 PheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHisCysAsp 162
Db 381 TTTAAGAGCTCTGTATAAATCGACGTACTTCCCGTTTGATTACCAAAACTGTACC 440
QY 163 MetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGlu 182
Db 441 ATGAAGTTCGGTTCCTGCTCCTACGATAAGGCGAAATCGATCTGCTGATCGGC--- 497
QY 183 AlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGlyMetPro 202
Db 498 TCTTCCATCAACTCAAGGACTATTGGGAGACGGCGAGTGGCCCATCATCAAGGCCCA 557
QY 203 GlyLysLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPhe 222
Db 558 GGCTCAAAACAGACATCAAGTACAACCTGCTCGAGGAGATCTACCCCGACATCACATAC 617
QY 223 ThrIleMetIleArgArgThrLeuTyrTyrPhePheAsnLeuIleValProCysVal 242
Db 618 TGCCTGTATACCCGGCGCTGCTGCTGTCTTACATCAATCAACTCATATCCCTGCCTG 677
QY 243 LeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeu 262
Db 678 CTCATCTCCTTCTCCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 737
QY 263 ThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThr 282

Qy 263 ThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThr 282
Db 738 ACGCTCGATCTCGGTCTCTCTCCCTGACGCTCTTCTCCCTGGATCACCGAGACC 797
Qy 283 LeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPhe 302
Db 798 ATCCCTTCCACCTCACTGCTCATCCCTTGATCGGGAGTACCTCTCTTCACTATGATT 857
Qy 303 MetValAlaSerSerValLeuThrValValValLeuLeuAsnTyrHisArgThrAla 322
Db 858 TTGTGACCTTGTCCATGCTCATCAGCTCTTTGTGCTCAACGTGCACTCAGAACTCCG 917
Qy 323 AspIleHisGluMetProGlnTyrIleLysSerValPheLeuGlnTyrLeuProTyrIle 342
Db 918 ACCACACACGATGCTCCACTTGGTCAAGCTGTCTTTGAACCTTCTCCCGAGGTC 977
Qy 343 LeuArgMetSerArgProGlyLysIleThrArgLysThrIleMetAsnThrArg 362
Db 978 ATGTTTATGACTAGGCCAACCCAGCAGGAGACGCCCAAGACGAGGAACCTTCTAC 1037
Qy 363 MetArgGluLeuGlu-----LeuLysGluArgSerSerLysSerLeuLeuAla 378
Db 1038 GGTGCTGAGCTCTCAACCTGAAGTCTGACGCGTGCAGACTCCAAAGCTGC----- 1091
Qy 379 AsnValLeuAspIleAspAspPheArgHisGlyProProProProAsnSerThrAla 398
Db 1092 -----AAGNAGGCTTACCCCTGCCAGATGGACATGT 1124
Qy 399 SerThrGlyAsnLeuGlyProGlyCysSerIlePheArg-----ThrAspPheArg 415
Db 1125 GGCTAT-----TGCCACCACCGTAGGGTAAAAATCTCAAAATTTTCAGT 1166
Qy 416 ArgSerPheValArgProSerThrMetGluAspValGlyGly-----GlyLeuGlySer 433
Db 1167 GCCAACTCACAAGACTCCAGTTCGAGTCTGTGTGATGCTGTGTGCTGCTGCT 1226
Qy 434 HisHieArgGluLeuHisLeuLeuLeuArgGluLeuGlnPheIleThrAlaArgMetLys 453
Db 1227 CTGTCCACCAAGTCAAGAGCCATCAAGTGTGAAGTATATGCTGAAATATGAAA 1286
Qy 454 LysAlaAspGluGluAlaGluLeuIleSerAspTyrLysPheAlaAlaMetValValAsp 473
Db 1287 GCACAGATGTAGCCAAAGAGATTAAGATGATGGAAGTATGTTGCCATGGTATGAT 1346
Qy 474 ArgPheCysLeuPheValPheThrPheThrIleIleAlaThrValAlaValLeuLeu 493
Db 1347 CGCATTTTCTCTGGGTTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1406

RESULT 7
AK080475 2010 bp mRNA linear HTC 19-SEP-2003
LOCUS Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
DEFINITION enriched library, clone:A730038F14 product:NEURONAL NICOTINIC
ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.
ACCESSION AK080475
VERSION AK080475.1 GI:26099260
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Carninci, P., Shibata Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
AUTHORS Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 10349636

MEDLINE
PUBMED
REFERENCE
AUTHORS20499374
11042159
3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS20530913
11076861
4

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS20530913
11076861
5

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS20530913
11076861
6

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS20530913
11076861
7

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES

source

1. .2010
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A730038F14"
/db_xref="MGI:2408016"
/db_xref="taxon:10090"
/clone="A730038F14"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="7 days neonate"
167. .1673

misc_feature

/note="NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3
SUBUNIT (SPTR|AAL58471, evidence: FASTY, 99.8%ID,
100%length, match=1497)
putative"

ORIGIN

anos

AY402878 1454 bp DNA linear GSS 15-DEC-2003
Mus musculus CHRN4 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY402878
AY402878.1 GI:39758861
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1. (bases 1 to 1454)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sminsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2. (bases 1 to 1454)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sminsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..1454
/organism="Mus musculus"

Qy	23	GlulysArgLeuLeuAsnAlaLeuLeu-----AlaAsnTyrAsnThrLeuGluArgPro	40
Db	24	GAGGAGAAGCTGATGGATGATCTCTCTGAACAAAACTCGCTACAAACAACTGATCGCCCA	83
Qy	41	ValAlaAsnGluSerGluProLeuGluValAlaArgPheGlyLeuThrLeuGlnIleIle	60
Db	84	GCCACCAGCTCTCCAGCTCATCTCCATCGCGCTGGAGCTATCACTGTCCAGCTCATC	143
Qy	61	AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsn	80
Db	144	AGTGTGAATGAGCGAGAACAGATCATGATGACCCACAGCATCTGGCTAAAAACAGGAATGGACT	203
Qy	81	AspTyrAsnLeuArgTrpTrpAsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThr	100
Db	204	GACTACCGCCCTGGCGCTGAAACAGCTCTCGCTATGAAGGGGTGAACATTCCTGAGGATTCCT	263
Qy	101	ProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp	120
Db	264	GCAAGCGAGCTGTGTCCTTGACATTTGTTGTGTACAAATGCCGATCGGACCTATGAG	323
Qy	121	GlyThrTyrGlnThrAsnValValAlaArgSerGlyGlySerCysLeuTyrValProPro	140
Db	324	GTGTCTGTCTACACCAAGTGTATCGTGTGGTTCCAACGGCAGCATCCAGTGGTGCCCTC	383
Qy	141	GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHis	160
Db	384	GCCATCTCAAGAGTGCCTGCAAGATAGAGGTGAAGACTTTCCCTTCGACCAGCAGAAC	443
Qy	161	CysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLys	180
Db	444	TGCACCTTCAATTTTCGTTCTCGTTCCTGGACCTATGACCACAGAGATTTGACATGTTCTTAA	503
Qy	181	AspGluAlaGlyLysLeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGly	200
Db	504	TCACCACA--GCATCATGATGATGACTTATCCCCAGCGGTGAATGGACATTTGTGGCC	560
Qy	201	MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGlu-----ProTyrVal	218
Db	561	CTCCCGGTTCGAAGG---ACGTGAAC-----CCACGAGACCCAGCTATGTG	605
Qy	219	AspValThrPheIleMetIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIle	238
Db	606	GACGTGACCTATGACTTTCATCATCAAGCGCAAGCCACTCTTCTACACCATCAATCTGATC	665
Qy	239	ValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSer	258
Db	666	ATCCCTTGTGTCTATCATCGTGTGGTATCTCTCTTACTCGCTCCGACTGT	725
Qy	259	GlyGluLysLeuThrLeuGlyValThrIleLeuSerLeuThrValPheLeuAsnLeu	278
Db	726	GGGGAGAAATGACGCTCTGCATCTCAGTGCTGTGCATCTCACATTCCTCTCTCTGCTGCTC	785
Qy	279	ValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsn	298
Db	786	ATCTCCAAGATCTGCTGCCCTCCACTCCCTTGACATTCCTCATTCATGGCAAGTACCTCTTG	845


```
Db      1046 ATCTCCAGATTGTGCTCCACCTCCCTCGACGTACCGTGTGGTGGAAAGTACCTCATG 1105
Qy      299 CysIleMetPheMetValAlaSerSerValValLeuThrValValValValLeuAsnTyrHis 318
Db      1106 TTCACCATGGTGTAGTACCTTCTCCATCGTCACTAGCGTGTGTGTCTCAATGTGCAC 1165
Qy      319 HisArgThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrp 338
Db      1166 CACCGTTCGCTACCGACACACATGCGCCCTCGGTCAAGTGTCTTCTCGGAGAAG 1225
Qy      339 LeuProTrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLysThrIleMet 358
Db      1226 CTGCCACCTCTCTCTCTCGACGACGACCGCTGTGCAGCTCAACGCTGC 1285
Qy      359 MetAsnThrArgMetArgGluLeuGluLysGluArgSerSerLysSerLeuLeuAla 378
Db      1286 TTGCGAAGCGCGGACGGAACGTGAG-----GGGGCAGGACACATTC--- 1330
Qy      379 AsnValLeuAspIleAspAspPheArgHisGlyProProProProAsnSerThrAla 398
Db      1331 -----TTCGCGAAGGCTCTGCGCTGGCTTTCACGGCTGACCGCTGACTGTC 1366
Qy      399 SerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgArgSerPhe 418
Db      1367 -----TTT 1369
Qy      419 ValArgProSerThrMetGluAspValGlyGly----- 429
Db      1370 GTCAATCTCTGCATCAATGCGAGGCTTGGCTGGGCTTTCACGGCTGAGCCGCTGCAGCC 1429
Qy      430 GlyLeuGlySerHisHisArgGluLeuHisLeuLeuLeuLeuLeuLeuLeuLeuLeu 445
Db      1430 GGCTGGGGCGCTCGATGGGCCCATGCGAGCTGCTGGCTCCGGGAGCTGTGCGAGGTGTA 1489
Qy      446 GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuLeuLeuLeuLeuLeu 465
Db      1490 CGCTTCATTCGGGACCATATGCGAAGTGAAGATGATGACACGAGTGTGAGGAGGATTGG 1549
Qy      466 LysPheAlaAlaMetValAlaAspArgPheCysLeuPheValPheThrLeuPheThrIle 485
Db      1550 AAATACGTTGCGATGGTGCATGCGACCGCTGTCTCTGGATCTTTGCTTGTCTGTGTC 1609
Qy      486 IleAlaThrValAlaValLeuLeu 493
Db      1610 TTTCGGACCATGCGATGTCCTG 1633

RESULT 11
LOCUS   AK051742
DEFINITION Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
enriched library, clone: D130070121 product: cholinergic receptor,
nicotinic, beta polypeptide 2 (neuronal), full insert sequence.
ACCESSION AK051742
VERSION   AK051742.1 GI:26342173
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
REFERENCE
AUTHORS   Carninci, P., and Hayashizaki, Y.
TITLE     High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED    10349636
REFERENCE
AUTHORS   Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED    11042159

REFERENCE
AUTHORS   Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hirozane, T.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED    11076861
REFERENCE
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium
TITLE     Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE     Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE
AUTHORS   Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murota, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
TITLE     Direct Submission
JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
FEATURES
source location/Qualifiers
1. .4046
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:D130070121"
/db_xref="MGI:2420100"
/db_xref="taxon:10090"
/clone="D130070121"
/tissue_type="spinal ganglion"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
209. .1714
/note="unnamed protein product; cholinergic receptor,
nicotinic, beta polypeptide 2 (neuronal) (MGD|MG1:87891,
GB|NM_009602, evidence: BLASTN, 99%, match=1498)
putative"
/codon_start=1
/protein_id="BAC34749.1"
/db_xref="GI:26342174"

CDS
```



```
/translation="WARCNSMALLPSPGLWLCGVLGDTDEERIVEHLDPFSRYNK
LIIRPATSGELIVTGLMVSQALISVHERBQITNTWLVLTQWBYSIATWKEPCEFN
KXVPLSGHILVDTLVXNADMGYSFVNAVSYDGSIFWLPALYKACEDIBK
HFPEQNCNTMKRSPYTDRIEIDLKVDASDLDTPEGSEWIDKIALPERNENPD
STVYDITDFIIRKRLPYTNLIPCVLITSLAIFVLPYDSCGEKMLTCLSVIAL
TVPLLSLFFVTSFVLPVGLKYMVLMTVLVTSIVTSCVPHHRRPSTYTHMAPV
KVVFLBKLPTLLFQOPHRCARQRLRRQREREGAGTLFREGPAADPCTQFVN
PASMGVLAGAFQABEAPAGLGRGSMGCGLRGAEVDVGRFIAHMRSDQSDVREDV
XAMVMDLRLEFATFVFCVFGTIGMFLQNYFQNTAATFLLHSKAPSSK"
```

Alignment Scores:

Pred. No.:	4,31e-95	Length:	4046
Score:	910.50	Matches:	185
Percent Similarity:	57.38%	Conservative:	95
Best Local Similarity:	37.91%	Mismatches:	153
Query Match:	34.49%	Indels:	55
DB:	11	Gaps:	9

US-09-303-232-6 (1-501) x AK051742 (1-4046)

Qy 23 GLYLysArgLeuLeuAsnAlaLeuLeu-----AlaAsnTyrAsnThrLeuGluArgPro 40
 Db 293 GAGGAGCGCGCTGGTGGAGCATCTCTTGGATCCCTTCGGCTATAACAAAGCTGATCCGTTCCA 352

QY 41 ValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIle 60
 ||| ||||| :: : : : :
Db 353 GCTACTAATGGGTCTGAGCTGGTGACTGTACAGCTCATGGTATCATTTGGCACGCTCATC 412

QY 61 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsn 80
413 AGTGTGCACAGCGGGAGCAGATCATGACACACACAGTCGTGGCTACCCGAGGATGGGAA 472
Db

Qy 81 AspTyrAsnLeuArgTirAsnAspSerGluTyrGlyValLysAspLeuArlleThr 100
 Db 473 GATTATCGCCTCACATGGAGCGCTGAGATTTCGACAAATATGAGAAAGTCCGACTCCCT 532

QY 101 ProAsnLysLeuTrpLysProAspValIleuMetTyrAsnSerAlaAspGluGlyPheAsp 120
Db 533 TCTAAGCACATCGCTCCAGAGTGCTTCTATACAAATGCTGACGCATGTACGAA 592

121 GlyThrTyrGlnThrAsnValValArgSerGlyGlySerCysLeuTyrValProPro 140
 593 GTCCTCTCTATTCCAAAGTCGTGGTCTCCCTATCATGGCAGCATCTTTTGGCTACCGCTT 652

Qy 141 GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHis 160
 Db 653 GCCATCTACAGAGCGCATGCAAGATTGAGTGGACACTTCCATTTTGCACGACAGAC 712

QY 161 CysaspMetLysPheGlySerTirThryAspGlyAsnGlnLeuAspLeuValLeuLys 180
 Db 713 TGCACCATGAAGTTCCGGTCTCTGGACCTACGACCGCACTGAGATTGACCTGGTGCTCAA 772

181 AspGluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTrpTyrIeuIleGly 200
773 AGCGATGTG---GCCAGCCTGGAGCACTTCACACCTAGTGGGAGTGGGACATCATCGCA 829

QY 201 MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGlu-----ProTyrVal 218
 ::::::::::::::::::::|||:::
 Db 830 CTGCCAGCCGACGCAACGAG-----AACCCAGACACTCCACCTACGTG 874

[illegible]

Qy ValProCysValLeulleSerSerMetAlaLeuLeuClyPheThrIeuProPaspser 258
:::|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Dbb ATCCCTCGGTACTCATCACTCGCTGGCCATCCTGTCCTTACCTGCCTCAGACTGT 994

259 GlyGluLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeu 278
 995 GGTAAAAGATGACATTTGATTCTGTGCTGCTGGCGCTACGGTGTTCCTGGTGTC 1054

QY	279	ValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsn	298
Db	1055	ATCTCCAAGATTGTGCCTCCACACCTCCCTCGACCTACCGTGGTGGGAAGTACTCATG	1114
QY	299	CysIleMetPheMetValAlaSerSerValValLeuThrValValValLeuAsnTyrHis	318
Db	1115	TTCACCATGGTGTAGTACACTTCTCATGTCTACTAGCGTGTGTGTGCTCAATGTGCAC	1174
QY	319	HisArgThrAlaAspIleHisGluMetProGlnThrIleLysSerValPheLeuGlnTrp	338
Db	1175	CACGTTCCGCTACACGACACCATGGCGCTTGGGTCAAGGTGCTTCTCTGGGAAG	1234
QY	339	LeuProTrpIleLeuArgMetSerArgProGlyLysIleThrArgLysThrIleMet	358
Db	1235	CTGCCACCTCTCTTCTTCGACAGCCACCGCTGTGCAGTCAACCCCTGGCG	1294
QY	359	MetAsnThrArgMetArgGluLeuGluLeuLysGluArgSerSerLysSerLeuLeuAla	378
Db	1295	TTGCGAAGCGCGACGCGGAACGTGAG-----GGGCAGCAGCACATATTC---	1339
QY	379	AsnValLeuAspIleAspAspPheArgHisGlyProProProAsnSerThrAla	398
Db	1340	-----TTCCGCGAAGGTCCTGCAGTCGACCCCATGTACCTGC	1375
QY	399	SerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgSerPhe	418
Db	1376	-----TTT	1378
QY	419	ValArgProSerThrMetGluAspValGlyGly-----	429
Db	1379	GTCAATCTGCATCAATGACGGCTTGGCTGGGGCTTCCAGGCTGAGCCCGCTGCAGCC	1438
QY	430	GlyLeuGlySerHisArgGluLeuHisLeuLeuLeuArgGlu-----Leu	445
Db	1439	GGTCTGGGCGCTCGATGGGCCCATGCAGCTGTGGCTCCGGGAAGCTGTGACGGGTGA	1498
QY	446	GlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGluLeuSerAspTrp	465
Db	1499	CGCTTATTTCGGACCATATGCAAGTGAAGATGATGACCAAGTGTGAGGGAGGATTGG	1558
QY	466	LysPheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle	485
Db	1559	AAATACGTTGCCATGTGATCGACCGCTGTCTCTGGAATCTTGTCTGTGTGTGTC	1618
QY	486	IleAlaThrValAlaValLeuLeu	493
Db	1619	TTTGGGACCAATTGGCATGTTCTGT	1642

RESULT 12

AL530299	AL530299	1201 bp	mRNA	linear	EST 23-MAY-2003
DEFINITION	AL530299 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens cDNA clone CS0DD007YP05 5-PRIME, mRNA sequence.				
ACCESSION	AL530299				
VERSION	AL530299.2	GI:31068132			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 1201)				
TITLE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
JOURNAL	Full-length cDNA libraries and normalization				
COMMENT	Unpublished (2001)				
	On Feb 13, 2001 this sequence version replaced the previous				

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7646.r For more information about this cluster, see
<http://www.genoscope.cns.fr/>

cgi-bin/cluster.cgi?seq=CS0DD007CH03QPI&cluster=7646.1. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DD007CH03QPI.

FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD007YP05"
/tissue="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="First strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
Pred. No.: 8.9e-96 Length: 1201
Score: 909.00 Matches: 181
Percent Similarity: 73.19% Conservatives: 51
Best Local Similarity: 57.10% Mismatches: 79
Query Match: 34.43% Indels: 8
DB: 9 Gaps: 2

US-09-303-232-6 (1-501) x AL530299 (1-1201)

QY 8 LeuAlaLeuLeuAla---LeuLeuProValSerGluGlnGlyProHisGluLysArgIleu 26
DB 152 CTGGCGCTGGCGGCTGCTCTCTGACGTCCTCTGCAAGCGAG-TWCCAGAGGAAGCTT 210
QY 27 LeuAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgProValAlaAsnGluSerGlu 46
DB 211 TACAAGGAGCTGGTCAAGAACTACAAATCCCTTGGAGAGCGCGTGGCCAAATGACGCGAA 270
QY 47 ProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleAspValAspGluLysAsn 66
DB 271 COACTACCGTCTACTTCTCTGAGCTCTCTGAGCATGACGATGATGAGAGAAC 330
QY 67 GlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 86
DB 331 CAAGTTTAAACCAACCAATTTGGTCTGCAAAATCTTGGACAGATCACTATTACAGTGG 390
QY 87 AsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThrProAsnLysLeuTrpLys 106
DB 391 AATGTGTGAGATATCCAGGGGTGAAGACTGTTCCTCCAGATGGCCAGATTTGGAAA 450
QY 107 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 126
DB 451 CCAGACATTTCTCTATACAGTGTCTGATGAGCGCTTGGACCCACATTCACACTAAC 510
QY 127 ValValValArgSerGlyGlySerCysLeuTyrValProProGlyIlePheLysSerThr 146
DB 511 GTGTGGTGAATTTCTTGGGATTTGCCAGTTCCTGATGAGCGCTTGGACCCACATTCAC 570
QY 147 CysLysMetAspIleAlaTrpPheProPheAspGlnHisCysAspMetLysPheGly 166
DB 571 TGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
QY 167 SerTyrThrTyrAspGlyAsnGlnLeuAspLeuValLeuLysAspGluAlaGlyGlyAsp 186
DB 631 TCCTGTCTTACGAGGCTGGTCTTGGATCTGACATGACGAG-----GCAGAT 681
QY 187 LeuSerAspPheLeuThrAsnGlyGluTyrThrLeuLeuGlyMetProGlyLysLysAsn 206
DB 682 ATCAGTGGCTATATCCCAATGAGAAATGGACCTAGTGGAAATCCCGGCAAGAGAGT 741
QY 207 ThrIleThrTyrAlaCysCysProGluProTyrValAspValThrPheThrIleMet 226
DB 742 GAAAGGTTCTATGAGTCTGCAAGAGCCCTACCCGATGTCACCTTCACAGTACCATG 801
QY 227 ArgArgArgThrLeuTyrTyrPhePheAsnLeuValProCysValLeuLeuSerSer 246

DB 802 CGCGCAGGAGCGCTTACTATGCGCTCAACCTGCTGATCCCTGTGTGCTCACTCGGCC 861
QY 247 MetAlaLeuLeuGlyPheThrLeuProAspSer-GlyGluLysLeuThrLeu-GlyV 266
DB 862 CTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 921
QY 266 alThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnV 286
DB 922 TAACAGTGTACTCTCTCTACCGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 981
QY 286 alSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAla 306
DB 982 CATCCGATTCGGTACCATTCGAGCCCGAGTCTTGGCAGCACCACCATCATCTGCTGG 1041
QY 306 erSerValValLeuThrValValValLeuAsnTyrHisHisArgThr 321
DB 1042 TTGG-GTGGTTGKRMGGTGTATCGTGTGCAGTACCACACACGACGACC 1087

RESULT 13
AY406230 1374 bp DNA linear GSS 15-DEC-2003
LOCUS
DEFINITION Homo sapiens CHRNA1 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION AY406230
VERSION AY406230.1 GI:39762204
KEYWORDS GSS.
SOURCE Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1374)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1374)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
1..1374
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1374
/gene="CHRNA1"
/locus_tag="HCM2488"

Alignment Scores:
Pred. No.: 3.3e-95 Length: 1374
Score: 905.00 Matches: 182
Percent Similarity: 56.02% Conservatives: 97
Best Local Similarity: 36.55% Mismatches: 149
Query Match: 34.28% Indels: 70
DB: 29 Gaps: 8

US-09-303-232-6 (1-501) x AY406230 (1-1374)

QY 4 MetLeuAlaLeuAlaLeuLeuAlaLeuLeuProValSerGluGlnGlyProHisGlu 23
DB 22 CTGCTCTTAGCCTTTGCTCAGCTGGCCTCGTCTGGCTCCGAA-----CATGAG 72

of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2940)

JOURNAL
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayaehida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, K., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES
source

1..2940
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9330165116"
/db_xref="MGI:2398424"
/db_xref="taxon:10090"
/clone="9330165116"
/sex="male"
/tissue type="dienecephalon"
/clone.lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="adult"
116..2005
/note="unnamed protein product; cholinergic receptor,
nicotinic, alpha polypeptide 4 (MGD|MGI:87888,
GB|NM_015730, evidence: BLASTN, 99%, match=1946)
putative"
/codon_start=1
/protein_id="BAC28638.1"
/db_xref="GI:26329799"

CDS

/translation="MEIGSGAGPPLLLPLLLGLTGLLPASSHETRAHAEERLLX
RLFGYKNSRPVANSIDVLRVFLGSLAQIDVDKRNQMTNWNVQKWHYKLRW
DPGDYNTSIRIPSELTWRDVIYNNADGPAVHTLKAHLYDGRVQNTPAIYK
SSCIDVTFPPDQONCKMGFSYDKADIDVSMHSRVQDLDFWESGEWIVDAVG
TYNRKVECCAEIYDPIYAIRPLPFTINLIPLLISCLTVLVPYLPSCGKEK
VTLGISLTSITVLLITLIPSTLVNPLIGEVLLFTMLFVTLIVITVFLNVHH
RSPHTHMPAWRRVFLDIVRELLFPMKRPSVVKNCRLIESMKMANAPRFPPEPES
EPGLIGDICNGLSFAPTFCKRMDTAVETQPCRSFHKVDFDLKTSVEKASPCPSFG
SCHPNSSGAPVLLKASLSVQHPVSEQAEGSIRCSRSIQYCVSQDGAASLPESK
TPGSPASLKTPTSPQSDTSPCKTCKEFPSPVITVLKAGGTAKPPQHLPLSPAL
TRAVEGVQYIADHLKAEDTDFSVKEDMKYVAMVIDRIELWVFIIIVCLLGTGVLFPWP
LAGMI"

ORIGIN

Alignment Scores:
Pred. No.: 5,92e-93 Length: 2940
Score: 890.50 Matches: 207
Percent Similarity: 48.05% Conservative: 88
Best local Similarity: 33.71% Mismatches: 195
Query Match: 33.73% Indels: 125
DB: 11 Gaps: 11

US-09-303-232-6 (1-501) x AK034228 (1-2940)

QY	3	ProMetLeuAlaAlaLeuAlaLeuLeu-----AlaLeuLeuProValSer	17
Db	143	CGCGCGTGTCTGCTCTCCGCGCTCCIGCTCTCTTAGGAGCCGCTCTTGGCTGTAGC	202
QY	18	-----GluGlnGlyProHis---GluLysArgLeuLeuAanAlaLeuLeuAan	33
Db	203	AGCCACATAGAGACCGCGGCCCATGCGGAGGAGCGGCTCTGAAGAGACTCTTCTCTGGC	262
QY	34	TyrAsnThrLeuGluArgProValAlaAsnGluSerGluProLeuGluValArgPheGly	53
Db	263	TACAACAAGTGTCTCGCGCCAGTAGCCAATATCTCAGATGTGGTCTCTGCGGTTTGGC	322
QY	54	LeuThrLeuGlnGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	73
Db	323	TTGTCGATTGCTCAGCTCATGTGATGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	382
QY	74	TrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyGly	93
Db	383	TGGGTGAAGCAGGAGTGGCATGACTACAACTGGCTGGGACCCCTGGTGTACTACAGAA	442
QY	94	ValLysAspLeuArgGlnLeuThrProAsnLysLeuTrpLysProAspValLeuMetTyrAsn	113
Db	443	GTCACTCCATCCGATCCCATCCCATCTGAATCATCTGGAGGCTGACATGCTCTTACAC	502
QY	114	SerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValValValValValVal	133
Db	503	AACCGGACGGGAGCTTTGAGTCAGTCACCCACCTAACCAAGCCACCTGTTTATGATGG	562
QY	134	SerCysLeuTyrValProProGlyIlePheLysSerThrCysLysMetAspIleAlaTrp	153
Db	563	CGTGTGAGTGGACACCCCGGCGCATCTATAAGAGCTCTCGACGATGACGATGACCT	622
QY	154	PheProPheAspAspGlnHisCysAspMetLysPheGlySerTrpThrTyrAspGlyAsn	173
Db	623	TTCCCTTCGACGACGAGACTTACCATGAAGTTTGGTCTGCTGAGCTACGACGACG	682
QY	174	GlnLeuAspLeuValLeuLysAspGluAlaGlyGlyAspLeuSerAspPheIleThrAsn	193
Db	683	AAGATTGACTGGTG---AGCATGCACAGCGCTGGACCACTGGACTTCTGGGAAAGT	739
QY	194	GlyGluTrpTyrLeuLeuGlyMetProGlyLysLysAsnThrIleThrTyrAlaCysCys	213
Db	740	GGGAGTGGGTCAATTGTGGATCGCGTGGGACCTACACACACGAGGAGTATGATGCTGT	799
QY	214	ProGluProTyrValAspValThrPheThrIleMetIleArgArgThrLeuTyrTyr	233
Db	800	GCCGAGATCTATCTCGACATCACCTACGCTTCATCCTCCGCGACTCCGCTGTCTAC	859
QY	234	PhePheAsnLeuLeuValProCysValLeuLeuSerSerMetAlaLeuLeuGlyPheThr	253
Db	860	ACCATCACTTATCATCCCGTGTCTCATCTCTCCCTCACCGTCTGTCTTCTAT	919
QY	254	LeuProProAspSerGlyGluLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThr	273
Db	920	CTGCCCTCGAGTGGCGGAGAGTCAAGTGTGATCGCTGTGCTGTCTCTCTCACC	979
QY	274	ValPheLeuAsnLeuValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeu	293
Db	980	GTCTTCTGCTGCTCATCCGAGATCATCCGCTCCCTCCCTGCTGTGTGTATCCGCTCATC	1039
QY	294	GlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValValLeuThrValVal	313
Db	1040	GCGAGTAGCTCTCTTCCACCATGATCTTGTCTCACCCTCTCCATGTCTACGCTCTTC	1099
QY	314	ValLeuAsnTyrHisHisArgThrAlaAspIleHisGluMetProGlnTrpIleLysSer	333
Db	1100	GTGCTCAATGTACACACCGCTCACACGACACACACACACACACACACACACACACAC	1159
QY	334	ValPheLeuGlnTrpLeuProTrpIleLeuArgMetSerArgProGly-----LysLys	351

Db	1160	GTCTTCTCGACATGTGCGCCGCTCTCTTTCATGAAGCGCCCATCTGGTGAAGAC	1219
Qy	352	IleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeuGluLeu	368
Db	1220	AACGCGGAGACTTATCGAATCCATGACACAGATGCGCAAGCGCCCTCGTTCTGGCCA	1279
Qy	369	LysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAsp	382
Db	1280	GAGCCTGAGAGTGAGCCCGGCATCTTGGGTGACATCTGCAACCAAGAGCCCTGTCACTGCC	1339
Qy	383	-----IleAsp-AspAspPheArgHisGlyProPr	392
Db	1340	CCAACTTCTGCACCCCATGGACACAGAGTTCAGACCCAGCCTACATGCAGGTACCC	1399
Qy	392	oProAsnSer	396
Db	1400	TCCACAAAGGTCCCTGACTTGAAGACATCAGAGGTTGAGAAGGCCAGTCCCTGTCCATCA	1459
Qy	396	-----	396
Db	1460	CTGGCTCTTGTCAACCACCAATAGCAGTGGGGCCAGTGTCTCATCAAGGCCAGGTCC	1519
Qy	397	-----ThrAlaSerThrGlyAs	402
Db	1520	CTGAGCGTCAGCATGTGCCAGTCCAGGAAGCAGCCGAGGCGACATCCGCTGCCGG	1579
Qy	402	nLeuGlyProGlyCysSerIlePheArgThrAspPheArgSerPheVal	419
Db	1580	TCTCGAGTATCCAGTACTGTGTTTC-CAAGATGGAGCTGCTTCCCTGACTGAGAGCAA	1638
Qy	419	-----	419
Db	1639	GCCACTGGCTCCCGACGCTCCCTGAAGACCCGTCATCCAGCTTCCAGTGTCCAGACCA	1698
Qy	420	-----ArgProSerThrMetGluAspVal	427
Db	1699	GACCTCTCCATGCMAATGCATGCATCAGGAACCATCTCCTGTGTCCCCCATCACTGTGCT	1758
Qy	428	-----GlyGlyGlyLeuGlySerHisHisArgGluLeuHisLeu	440
Db	1759	CAAGGCTGGAGGCCAACCAAGACCTCCCAACACCTCCCTGTCCAGGCCCTGACACG	1818
Qy	441	IleLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAlaAspGluAlaG	460
Db	1819	GGCAGTAGAAGCGTCCAGTACATTTGCAGACCACTCAAGGACGAGACACAGACTTCTC	1878
Qy	460	nLeuIleSerAspTrpLysPheAlaAlaMetValValAspPheCysLeuPheValPh	480
Db	1879	GGTGAAGAGGACTGGAAATACGTGGCCATGTGTCATGACCGAATCTTCTCTGGATGTT	1938
Qy	480	eThrLeuPheThrIleIleAlaThrValAlaValLeuLeu	493
Db	1939	CATCATTTCTCGCTTCTGGGCACCTGTGGGACTCTTCTG	1978
RESULT 15			
AK083157			
LOCUS	AK083157	3230 bp	mRNA
DEFINITION	Mus musculus adult male hippocampus cDNA, RIKEN full-length enriched library, clone: C630019M18 product: cholinergic receptor, nicotinic, alpha polypeptide 4, full insert sequence.		
ACCESSION	AK083157		
VERSION	AK083157.1	GI:26350296	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		

```
/dev_stage="adult"
36..1925
/notes=unnamed protein product; cholinergic receptor,
nicotinic, alpha polypeptide 4 (MGD|GI:87888,
GB|NM_015730, evidence: BLASTN, 99%, match=1946)
putative"
/codon_start=1
/protein_id="BAC38788.1"
/db_xref="GI:26350297"
/tranlation="MEIGSGAPPPLLLPLLLLTGLLPASSHITRAHAERLIK
RLFSYKMKRPANISDVLRFLGLIAQLIDVDKRNQMTTNVWKQEMHDKLRM
DPGDIYNTSIRIPSELIWRPDI VLYNNADGFVTHLTKAHLFYDGRVQWTPPAIYK
SSGDIYVTFPDQOQNTKMFSGWYDKAKIDVMSHRVDQLDFWESGWIVDVGK
TYNRKYECACAEIYPDITYAFIIRRLPLFTINLIIPCLISLTVLVFVLPSCGK
VTLGICSLSTVTLLEIIPSTSLVPLIGLFTWLFVTLSTVIVFVNLVNH
RSPRTHMPAQRVRFIDIVPRLLFMKRPVVDKNCRLJESMEKMANARFPPEPS
EPGLGIDICNGSPATFCNRMNTAVETQTCRSPKHPDLDLTKESKAPCPSPG
SCHPNSSGAPVLKARLSVQHVPSSEAEAGSIRCSRISIQYCSQDGAASLTSG
FTSPASLKRTPSOLPVSQDTPCKCTCKEFPSPVITVLKAGTKAPQHLPLSPAL
TRAVEGVQYIADHLKABEDTDFSVKEDWKYVAMVIDRIFLWMFIIVCLLGTGLFLPWP
LAGMI"
ORIGIN
Alignment Scores:
Pred. NO.: 6.9e-93 Length: 3230
Score: 890.50 Matches: 207
Percent Similarity: 48.05% Conservative: 88
Best Local Similarity: 33.71% Mismatches: 195
Query Match: 33.73% Indels: 125
DB: 11 Gaps: 11
US-09-303-232-6 (1-501) x AK083157 (1-3230)
QY 3 ProMetLeuAlaLeuLeuLeu-----AlaLeuLeuProValSer 17
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 63 CCGCGGTGCTGCTCTCGCGCTCTCTGCTCTTAGGACCGGCTCTTTCCTGCTAGC 122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18 -----GluGlnGlyProHis---GluLysArgLeuLeuAsnAlaLeuAsn 33
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 123 AGCCACATAGAGACCGCGGCCCTGCGGAGGAGCGGCTCTTGAAGAGACTCTCTCTGCG 182
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 34 TyrAsnThrLeuGluArgProValAlaAsnGluSerGluProLeuGluValArgPheGly 53
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 183 TACAAAGTGGTCTCGCGCAGTAGCAATATCTCAGATGTGGTCTTGTCTGCGTTTGGC 242
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 54 LeuThrLeuGlnGlnIleLeuAspValAspGluLysAsnGlnLeuLeuIleThrAsnIle 73
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 243 TTGTGATTGCTCAGCTCATTGATGTGGATGAGAGAAAACCATGATGACGACCAACGTG 302
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 74 TrpLeuSerLeuGluTrpAsnAspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyGly 93
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 303 TGGGTGAAGCAGGAGTGGCATGACTACAAACTGGCTGGGACCTCTGTGACTACGAGAT 362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 94 ValLysAspLeuArgIleThrProAsnLysLeuTrpLysProAspValLeuMetTyrAsn 113
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 363 GTACCTCCATCGCATCCCATCTGAACTCATCTGGAGGCTGACATCGCTCTCAAC 422
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 114 SerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValArgSerGlyGly 133
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 423 AACGGGACGGGACTTTGAGTCAGTCACCCACTAACCAAGCCACCTGTCTATGATGGG 482
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 134 SerCysLeuTyrValProGlyIlePheLysSerThrCysLysMetAspIleAlaTrp 153
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 483 CGTGTGAGTGGACACCCCGGCGCATCTATAAGAGCTCTCTGACATCGACGTCACTTC 542
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 154 PheProPheAspGlnHisCysAspMetLysPheGlySerTrpThrTyrAspGlyAsn 173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 543 TTCCCCITCGACACAGAGACTGACCATGAAGTTTGGGTCCTGGACCTACGACCAAGGCC 602
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 174 GlnLeuAspLeuValLeuLysAspGluAlaGlyAspLeuSerAspPheIleThrAsn 193
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 603 AAGATTGACTTGGTG---AGCATGCACAGCGCTGTGGACCAACTGGACTTCTGGGAAGT 659
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
194 GlyGluTrpTyrLeuIleGlyMetProGlyLysLysAsnThrIleThrTyrAlaCysCys 213
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
660 GGGGAGTGGGTCAATGTGGATGGCGTGGGCACCTACAACACGAGAGTATGAATGCTGT 719
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
214 ProGluProTyrValAspValThrPheThrIleMetIleArgArgArgThrLeuTyrTyr 233
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
720 GCCGAGATCTATCTCTGATCACCCTACGCTTCATCATCCGCCACTGCCACTGCTTAC 779
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234 PhePheAsnLeuIleValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThr 253
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
780 ACCATCAACCTTATCATCCCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAT 839
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
254 LeuProProAspSerGlyGluLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThr 273
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
840 CTGCCCTCGAGTGGCGGAGAGGTCACTGTGTCATCTCGGTGCTGCTTCTCTCTAC 899
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
274 ValPheLeuAsnLeuValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeu 293
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
900 GTCTTCTGCTGTCTCATCACCAGATCATCCCGTCCACTCGCTGGTTCATCCCGTCTATC 959
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
294 GlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValLeuThrValVal 313
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
960 GCGGAGTACCTGCTCTTCCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1019
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
314 ValLeuAsnTyrHisHisArgThrAlaAspIleHisGluMetProGlnTrpIleLysSer 333
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1020 GTGCTCAATGTACACACCGCTCACCAGGCACACACACCATGCGCGCTGGTGGCGAGA 1079
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
334 ValPheLeuGlnTrpLeuProTyrIleLeuArgMetSerArgProGly-----LysLys 351
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1080 GTCTTCTGAGCATTTGTGCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1139
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
352 IleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeuGluLeu----- 368
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1140 AACTGCCGAGACTTATCGAATCCATGCACAGATGGCCAGCCCTCGTTCTTCTGGCCA 1199
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
369 LysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAsp----- 382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1200 GAGCTTGAGAGTGGCGCGCATCTTGGGTGACATCTTGCAACCAAGCCCTGTCACCTGCC 1259
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
383 -----IleAsp-AspAspPheArgHisGlyProPr 392
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1260 CCAACTTCTGCAACCGCATGGACACAGCAGTCGAGCCACCGCTATCATGAGGTCAACC 1319
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
392 oProAsnSer----- 396
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1320 TCCACAAAGTCCCTGACTTGAAGACATCAGAGTTGAGAGGCCAGTCCCTGTCTCATCA 1379
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
396 ----- 396
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1380 CCTGGCTCTTGTCAACCAACCAATAGCAGTGGGCCCCAGTGTCTATCAAAAGCCAGGTCC 1439
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
397 -----ThrAlaSerThrGlyAs 402
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1440 CTGAGCGTCCAGCATGTGCCAGTCCAGGAAGACAGCCGAGGCGCATCTCGCTGCGGG 1499
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
402 nLeuGlyProGlyCysSerIlePheArgThrAspPheArgArgSerPheVal----- 419
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1500 TCTCGAGTATCAGTACTGTGTTCC-CAAGATGGAGCTGCTTCCCTGACTGAGAGCAA 1558
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
419 ----- 419
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1559 GCCCACTGGCTCCCGCCCTCCCTGAAGACCCGTCATCCAGCTTCCAGTGTGAGACCA 1618
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
420 -----ArgProSerThrMetGluAspVal----- 427
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1619 GACCTCTCCATGCAATGACATGCAAGGAACCATCTCTGTGTGCTCCCTCATCTGTGCT 1678
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
428 ----GlyGlyGlyLysSerHisArgGluLeuHisLeu----- 440
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1679 CAAGGTGGAGGACCAAGACCTCCCCAACACCTGCCCTCTCACCAGCCCTGACACG 1738
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
441 -IleLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAlaAspGluAlaGlu 460
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 1739 GGCAGTAGAAGGCGTCCAGTACATTGCAGACCCTCAAGGCGAAGACACAGACTTCTC 1798
Qy 460 uLeuileSerAspTrrPheAlaMetValValAspArgPheCysLeuPheValPh 480
Db 1799 GGTGAAGGAGGACTGGAAATACGTGGCCATGGTCAATTGACAGAAATCTTCTCTGGATGTT 1858
Qy 480 eThrLeuPheThrIleAlaThrValAlaValLeuLeu 493
Db 1859 CATCATTTGTGCTTCTGGGCACTGTGGGACTCTTTCCTG 1898
```

Search completed: May 9, 2004, 04:21:03
Job time : 3532.21 secs

This Page Blank (uspto)